









# CHEMICAL REVIEWS

GERALD WENDT, Editor  
LOUISE KELLEY, Assistant Editor

VOLUME 15

PUBLISHED BI-MONTHLY FOR  
THE AMERICAN CHEMICAL SOCIETY  
BY  
THE WILLIAMS & WILKINS COMPANY  
Baltimore, U. S. A.  
1934



## CONTENTS

### NUMBER 1, AUGUST, 1934

Active Iron. I. The Structure of Werner Complexes of Metals of the Iron Family in Relation to the Chemical, Physical, and Biological Behavior of Their Central Metal Atoms. OSKAR BAUDISCH AND LARS A. WELO.....	1
Active Iron. II. Relationships among the Oxide Hydrates and Oxides of Iron and Some of Their Properties. LARS A. WELO AND OSKAR BAUDISCH.....	45
The Chemical Nature of Enzymes. HENRY TAUBER.....	99
Colloid Developments in Synthetic Resins. H. L. BENDER, H. F. WAKEFIELD, AND H. A. HOFFMAN.....	123

### NUMBER 2, OCTOBER, 1934

Changes in Crystal Structure and Their Influence on the Reactivity and Catalytic Effect of Solids. J. ARVID HEDVALL.....	139
Hydrazoic Acid and Its Inorganic Derivatives. L. F. AUDRIETH.....	169
Principles of an Electronic Theory of Organic Reactions. CHRISTOPHER K. INGOLD.....	225

### NUMBER 3, DECEMBER, 1934

The Haloform Reaction. REYNOLD C. FUSON AND BENTON A. BULL..	275
The Chemistry of the Bile Acids and Related Substances. HARRY SOBOTKA.....	311
The Physical Chemistry of Foams. SOPHIA BERKMAN AND GUSTAV EGLOFF.....	377
$\beta, \beta'$ -Dichloroethyl Sulfide (Mustard Gas). KIRBY E. JACKSON.....	425



## ACTIVE IRON. I

### THE STRUCTURE OF WERNER COMPLEXES OF METALS OF THE IRON FAMILY IN RELATION TO THE CHEMICAL, PHYSICAL, AND BIOLOGICAL BEHAVIOR OF THEIR CENTRAL METAL ATOMS

OSKAR BAUDISCH AND LARS A. WELO

*Sterling Chemistry and Sloane Physics Laboratories, Yale University, New  
Haven, Connecticut*

*Received April 3, 1934*

Gestaltung, Umgestaltung,  
Des ewigen Sinnes ewige Unterhaltung.—Faust II. Teil.

#### I. INTRODUCTION

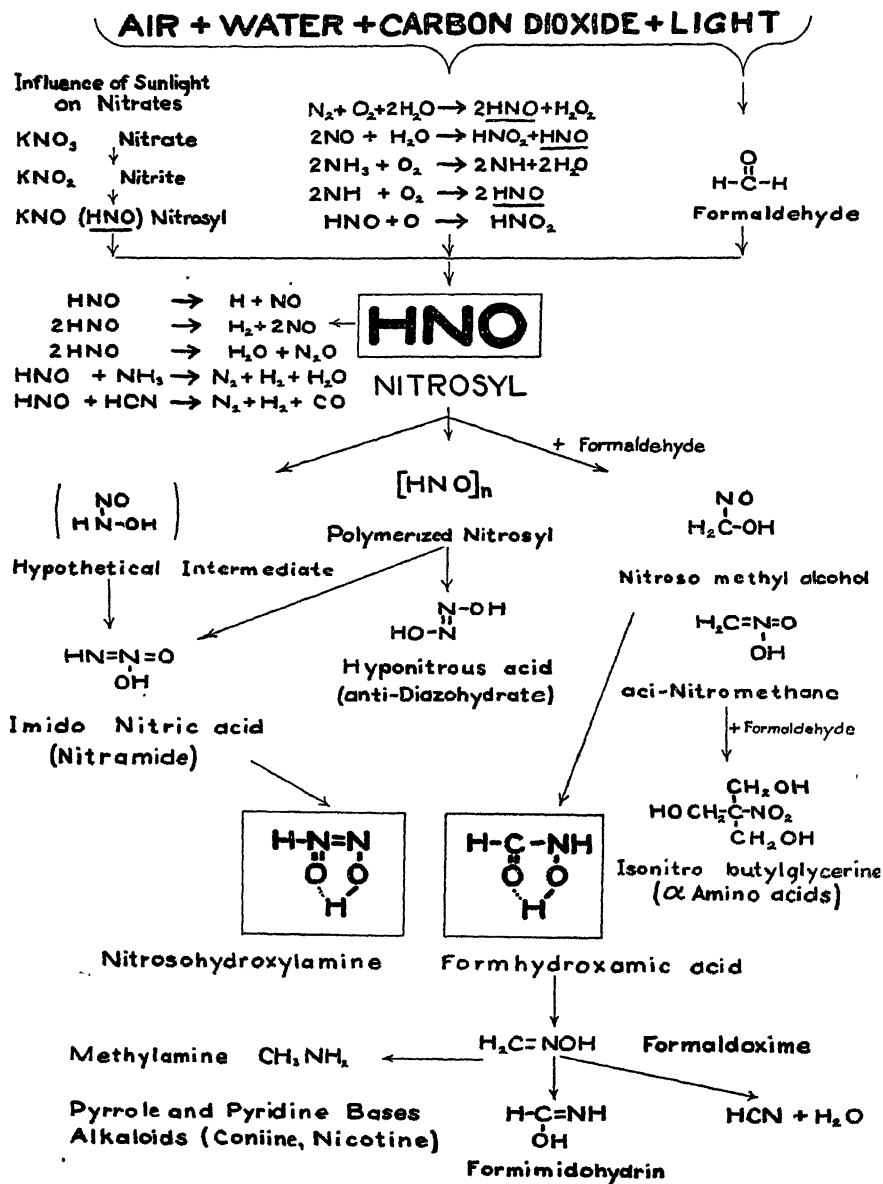
In the present review we shall include the work carried out to investigate in detail the Werner complex salts (84) with organic addenda and a central metal ion, especially the so-called inner complex salts (11). The compounds and complex-forming groups which we have used for this purpose were not chosen freely but were retained from the study of fundamental photochemical processes, which were carried out to synthesize organic compounds from inorganic compounds by photochemical methods. In this, we were especially interested in the nitrogen cycle; but the simplest inorganic compounds which are concerned in the building up of organic matter—nitrogen, ammonia, and nitrates—were chosen. As the source of carbon we used formaldehyde chiefly, without consideration of the question of its photochemical formation from carbon dioxide.

In table 1 the results of our photochemical investigations are summarized.<sup>1</sup> Nitrosyl, HNO, functions as the most important compound, as far as the nitrogen in the synthesis of simple carbon- and nitrogen-containing materials is concerned. It acts as a branching-point, leading to several series of compounds. The most important organic compound synthesized by light, which we have isolated and analyzed as an iron and copper complex copper salt, is formhydroxamic acid. From it by the action of light we obtained its labile isomer, acinitromethane, as well as formaldoxime. Formaldoxime in light undergoes a Beckmann rearrangement and gives formimidohydrin or formamide. We note in the table the

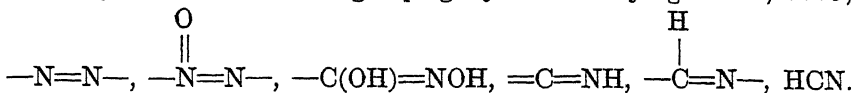
<sup>1</sup> A separate bibliography of the publications on which this table is based appears at the end of this article.

TABLE I

Formation of complex-forming groups and of organic compounds from inorganic by the influence of light



following characteristic atom groupings synthesized by light: NO, HNO,



We obtained compounds and atom groups all of which are capable of complex formation with iron and copper, and of which a great number of characteristic iron and copper complex salts have been described in the literature of organic chemistry (72, 80). This fact led us to include in our investigations the metal complex salts, especially the iron and copper salts, of these characteristic complex-forming groups. It is of especial significance that simply by exposing mixtures of alkali nitrites and formaldehyde to sunlight, or by irradiation of aldoximes in the presence of formaldehyde, highly polymerized, dark colored, syrupy substances were obtained. The presence of the pyrrole and also the pyridine ring was established with certainty by characteristic color reactions and by the isolation of nicotine. This fact is of interest from the point of view of plant physiology. The great affinity of the pyridine ring, and in general of nearly all nitrogen-containing rings for the iron atom, is not, we believe, conditioned only by the dipole nature of the pyridine, but is to be attributed chiefly to the special chemical and physical character of nitrogen linked in a ring. We shall later describe experiments dealing with this point.

The particular affinity of nitrogen compounds for iron, especially of the nitrogen-oxygen compounds, is noteworthy even in the simplest nitroso compound, nitric oxide, NO. All organic nitroso compounds, furthermore, the hydroxylamines and the oximes, show characteristic affinity for iron, which we will treat in detail in this review. In compounds of nitrogen with carbon, e.g., in hydrocyanic acid and in complex organic compounds which contain a  $=\text{C}=\text{NH}$ , or a  $-\text{CH}=\text{N}-$  group, the selective affinity for divalent iron especially is exceptionally notable. We were able by this means to split, by oxidation, members of the purine, pyrimidine, and hydantoin series (50).

Early in our investigations we emphasized the fact that organic addenda which are drawn into the inner sphere of a metal atom are activated and thereby become more easily available for oxidative cleavage or chemical replacement (14, 15, 21, 64, 65). The linkage to the central metal ion—in the special cases considered here, to the iron ion or atom—is dependent on the chemical composition of the linked organic compounds. The central metal ions together with the addenda form “complex compounds” which possess a characteristic geometric crystal structure. The atoms or molecules attracted into the sphere of the central atom exert a so-called “field force,” and substances with similar fields may replace them isomorphously in the crystal lattice, i.e., such substances are “perceived” as



identical by the growing crystal. The field similarity depends essentially on the number and the arrangement of the outer electrons in the atoms or molecules. In trying to trace the appearance of isomorphism in a broader sense back to a fundamental lattice constant, Langmuir (55, 56; see also 62) was the first to apply atomic theory to the problem. Langmuir expressed the thought that the similarity of the crystallographic ratios in a series of chemically unrelated substances, to which Barker (7; also 44) had first called attention, depended on *isosterism*. This means that in the unit cells of the substances compared, the outer electrons must be the same in number and arrangement.

It is of particular interest that the compounds and atom groupings synthesized by us, using light energy, form such isosteric groups, as may be seen in the following:

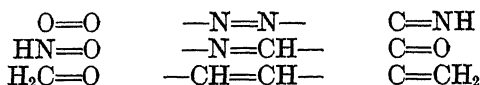


TABLE 2

*Diamagnetic susceptibility values for different atomic groups*

GROUPS	DIAMAGNETIC SUSCEPTIBILITY VALUES
O= (aldehyde)	+17.5
O= (nitroso)	+79.5
C=C	+54.5
N=N—	+18.5
C=N—	+81.5

Furthermore, it is interesting that the groups given above show a paramagnetic behavior<sup>2</sup> and thereby reveal a certain unsaturated condition in the molecule (table 2, from Pascal (63)). That there is also in the metal complex salts a particular state of affairs is shown by the magnetic measurements of metal salts, which have been extensively investigated by Welo (81).

The following types of complexes can be distinguished:

- (1) Those which show the magnetism of the central ion. Examples are benzohydroxamic acid iron salt and the iron salt of nitrosophenylhydroxylamine (cupferron iron). These point to the concept that between the central ion and the addendum only ions or dipoles are attracted. The symmetry relations which determine the magnetism are not disturbed here, and one may suppose that no inner fusion of the electron structure of the addenda with the central ion has taken place.

<sup>2</sup> See however, E. C. Stoner, *Magnetism and Atomic Structure*, p. 36, E. P. Dutton and Co., for another interpretation of this effect.

- (2) Those complexes which show a fundamentally different magnetism than that of the central atom or ion. Here it is presumably the case that a part of the electron structure of the addendum is attracted so close to the central ion that a complete rearrangement has occurred.

The complexes of the first class were called by W. Biltz (33) "normal;" those of the second class were called by P. Ray (69) "(penetration) complexes." Table 3, by W. Klemm, H. Jacobi, and W. Tilk (52), shows that there is here a sharp distinction. The normal complexes possess the magnetic moment of the  $\text{Fe}^{++}$  or  $\text{Fe}^{+++}$  ions; the penetration complexes are in part diamagnetic and in part are paramagnetic to be sure, but their

TABLE 3

*Weiss magneton numbers for some complex compounds of iron (to nearest unit)*

$\text{Fe}^{++}$ IONS	MAGNETON NUMBERS	$\text{Fe}^{+++}$ IONS	MAGNETON NUMBERS
A. Normal complexes			
$[\text{Fe}(\text{H}_2\text{O})_4]\text{Cl}_2$ .....	26	$\text{Na}_3[\text{FeF}_6]$ .....	29
$[\text{Fe}(\text{H}_2\text{O})_6]\text{SO}_4 \cdot \text{H}_2\text{O}$ .....	26	$\text{Na}_3[\text{Fe}(\text{SO}_4)_3]$ .....	29
$\text{FeCl}_2$ .....	27	$\text{Fe}_2(\text{SO}_4)_3$ .....	29
$\text{FeSO}_4$ .....	26	$\text{FePO}_4$ .....	28
B. Penetration complexes*			
$\text{K}_4[\text{Fe}(\text{NC})_6]$ .....	0	$[\text{Fe}(\text{CO})_5]$ .....	0
$\text{K}_3[\text{Fe}(\text{NC})_6]$ .....	12	$[\text{Fe}_2(\text{CO})_9]$ .....	0
$\text{Na}_3[\text{Fe}(\text{NC})_5 \cdot \text{NH}_3]$ .....	0	$[\text{Fe}_3(\text{CO})_{12}]$ .....	0
$\text{Na}_3[\text{Fe}(\text{NC})_5\text{NO}]$ .....	0	$[\text{Fe}(\text{dipyridyl})_3]\text{Br}_2$ .....	0
$\text{Na}_3[\text{Fe}(\text{NC})_5\text{NO}_2]$ .....	11	$\text{Na}_3\text{Fe}(\text{NC})_5\text{SO}_3 + 8.8 \text{ H}_2\text{O}$ ...	0

\* To differentiate between the valences in these compounds would have no meaning.

magnetic moments are very much smaller than those of the simple ions. Therefore, in this case, one can with certainty assign a substance to one or the other of the two classes on the basis of the magnetic measurements. It is therefore shown by this table that the magnetic behavior of the complex is the most definite criterion for the differentiation of the two different types of complexes.

According to W. Biltz (34; also 67), the theory of the diamagnetic cobalt salts requires a close linking of the nitrogen with the central atom, exhibited by the sharing of two of the nitrogen electrons by the two atoms, similarly to the non-polar bond between carbon and its addenda which we assume to exist in organic substances. X-ray study of luteo iodide yields a direct measurement of the shortest cobalt-nitrogen linkage. As a consequence

of this diminution of the cobalt-nitrogen distance, there is a possibility of assigning a smaller space to the ammonia molecule. This possibility is realized in the luteo salts, all of which have equal cobalt-nitrogen spacing. Chemical results are in agreement with this structural concept, which is equivalent to that of organic molecules.

Exactly as in the case of cobalt, the complex compounds of iron either retain the moments of the  $\text{Fe}^{++}$  or  $\text{Fe}^{+++}$  ions—in which case the central ion and the addendum remain apart—or they undergo a profound change in which the iron and the added constituent couple to form a single system. There are, for example, complex iron compounds which are more weakly paramagnetic the firmer the complex bond of the iron, in other words, the higher the degree of complexity. According to a theory introduced by Welo and Baudisch (82), it is assumed that the paramagnetism, which is due to unpaired electrons of the central atom, is reduced by the linking on of complex addenda so that new electron pairs are formed with consequent lowering of the number of unpaired electrons. Ladenburg, like Klemm (53) and his coworkers, assumes that, by the linking of the unpaired outer electrons of the iron with electrons of the addenda, an "intermediate layer" is produced about the central atom, and that diamagnetism is obtained whenever this intermediate layer contains sixteen or eighteen electrons, the latter being a rare gas configuration.

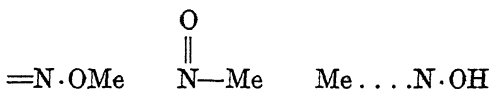
According to experiments of Cambi and Szegő (38), the entrance of electrons into this intermediate layer, the third level of iron, occurs only with addenda with weak dipole properties. On the other hand, addenda with strong dipole properties are deformed by the central ion without leading to the formation of the intermediate layer. One finds, therefore, in this class of complexes the same paramagnetic moments as in the simple ionic compounds of iron.

We have found that iron compounds which are chemically very complex—i.e., form a stable complex, as for example the iron cupferron salt—retain the high paramagnetism of the  $\text{Fe}^{+++}$  ion. Iron benzohydroxamate behaves similarly. In this case the chemical behavior contradicts the magnetic behavior. In this connection, interesting and important work has been done recently by Felix Haurowitz and Hubert Kittel (46; see also 19), who have investigated the magnetic properties of some hemoglobin derivatives. From the magnetic properties of the three hemins investigated, it is demonstrated that here also the iron atom has retained the high paramagnetism of the  $\text{Fe}^{+++}$  ion in combination in the molecule and that, therefore, no stable complex linkage in the sense of the previously discussed observations of Ladenburg, Klemm, and Cambi has been formed. Chemically, the iron in hemin is very firmly bound, for hemins can be treated with boiling hydrochloric acid or alkalies without removal of the iron. Hauro-

witz and Kittel consider that the iron in hemins is indeed linked by only weak forces, but that it is protected against the dissociating action of water and the influence of acids and alkalis by the characteristic structure of the porphyrin nucleus.

Our work is concerned with the clearing up of the special type of linkage of metals, particularly iron and copper, with organic nitrogen-containing addenda and we have sought to relate the chemical and biological functions of these metals with their special complex linkage and with their magnetic properties. With this in view, Welo (81) has measured magnetically a great number of complex metal salts.

For our purpose, we have used especially nitroso and oxime compounds, since it is known that the NO and HNO groups have a great influence on the magnetic susceptibility of iron ions, and since the manner of linkage of the HNO or of the oxime group to the metal appears to be determining for its complex-chemical behavior. The following simple metal compounds with the =NOH group were found:



It is of determining significance for the chemical and also for the magnetic character of the oxime complex, whether the metal is linked by primary valence to *oxygen* or to *nitrogen*. Moreover, linkage through secondary valences to nitrogen of the  $\dots \text{NOH}$  group is quite characteristic of certain complexes. However, there can also be combinations of all of these types of linkage, and in such cases it is difficult to determine the fine structure of such complexes. The character of the metal (8) itself, i.e., its position in the periodic system, is also very important in governing the chemical and physical properties of its complexes, as we shall see later.

## II. GENERAL DISCUSSION OF COÖRDINATION VALENCY AND COÖRDINATION COMPOUNDS

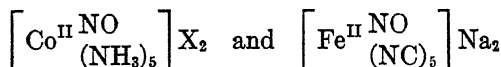
Werner (84) developed the concept of coördination valency as applied to compounds which are formed by addition of ammonia to certain heavy metal salts, the metal ammonia compounds. The stability of the ammonia complexes was thoroughly studied by Ephraim (42) and by W. Biltz (32) and collaborators. Their results showed that the stability of these compounds was increased as the charge on the central ion increased and as its size decreased. The same rules apply to the linking of water of crystallization; with the increase in size of the cation the tendency for the formation of hydrates diminishes. Thus calcium chloride crystallizes with six, and barium chloride crystallizes with two, molecules of water.

Elements do not possess the ability to exert coördination valence in all compounds. Thus, for example, nitrogen can develop a coördination valence in ammonia and all of its derivatives but not in ammonium compounds. The nitrogen of nitrites exerts coördination valence toward iron, but that of nitrates does not.

Nitric oxide (NO) exerts coördination valence toward iron, e.g., in the compound<sup>3</sup>



Those compounds are decidedly more stable in which the NO is linked to a metal atom along with other types of neutral fragments, such as  $\text{NH}_3$  or CN:



Oxygen and also sulfur can exert coördination valence in certain compounds (metal oxides, ketones, aldehydes, carbonyl oxygen atoms, thio-urea, etc.).

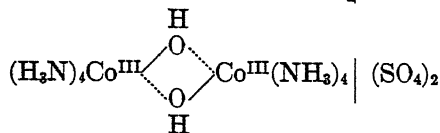
According to the Werner theory, coördination compounds can be divided into three classes.

The *first class* consists of those compounds which arise by coördinative linkage to atoms, the principal valences of which are partially or completely saturated. The first group of this class is made up of those in which the union has taken place between two or more molecules by coördination valences, e.g.:



In this compound the phosphorus is so firmly bound by coördination valences to the platinum that water will hydrolyze the chlorine atoms and form platinum dichloro-diphosphorous acid.

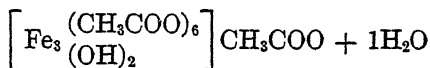
Polynuclear compounds, which play a part in the aging of metal hydrates, are formed among other ways by coördinative linking to the oxygen atoms of hydroxyl groups, e.g.:



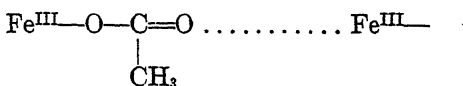
Octammino- $\mu$ -diol-dicobaltisulfate

<sup>3</sup> NO as a monovalent positive radical is called nitrosyl, while as a neutral fragment, as in the above compound, it is called the nitroso group.

Weinland (80) has prepared a great number of metal salts of organic acids, e.g., the deep red ferric acetate is the acetate of a polynuclear tri-ferri-hexaacetato-dihydroxo cation with the formula

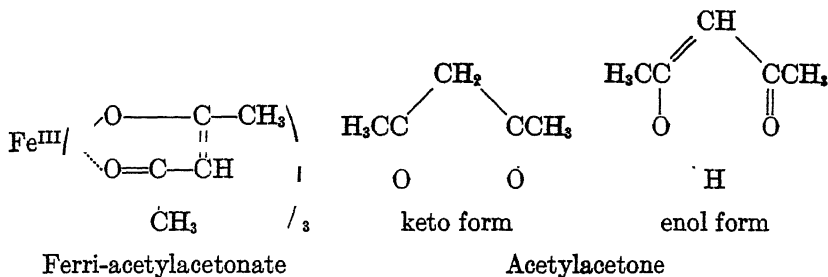


The linkage of the iron atom takes place through coördination valence which is located on the carbonyl oxygen atom of the carbonyl group.



Welo (81) has measured a large series of the Weinland salts magnetically and found that in almost every case they show the normal magnetic moment of the  $\text{Fe}^{+++}$  ion.

The second group is made up of the compounds characterized by the operation within a single molecule of both principal and subsidiary valences. The salts of this type belong to the *inner complex salts* discussed earlier. An example is:



This iron compound represents chemically a highly complex compound, in which the properties of the ferric ion are completely masked. However Welo's measurement of it shows it to be derived from the iron ionogen in a magnetic sense, for it shows 29 Weiss magnetons, as does the iron in ordinary ionogenic salts.

The blood pigment and chlorophyll belong to the inner complexes. It is interesting physiologically that the iron in hemin, according to the most recent magnetic measurements of Haurowitz and Kittel (46), has retained the high paramagnetism of the  $\text{Fe}(\text{III})$  ion. It completely resembles in this respect the ferri-acetylacetonate or cupferron iron. Magnetically they are ionogenic salts; chemically they appear to be strong complex salts in which the chemical properties of the central metal have been lost.

The third group are the organic molecular compounds, such as the

quinhydrone, choleic acid, and addition compounds of polysaccharides and proteins.

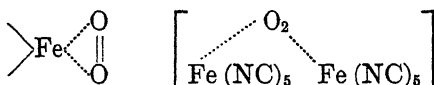
The *second class* of coördination compounds consists of those in which an uncombined atom is linked to one the principal valence of which is satisfied, as in the yellow silver subfluoride.



The ammonia compounds of the alkali and alkaline earth metals are examples of compounds of elements with molecules; thus the lustrous red compound (35):



Further illustrations are oxy- and carbon monoxy-hemoglobin, and peroxo-ferro-aquo salt (30):



The *third class* of coördination compounds is formed by the amalgams, such as  $\text{KHg}_{12}$  and  $\text{CaHg}_{12}$ .

The number of coördination valences which an element possesses is always one in the case of the non-metals of low atomic weight, as nitrogen, oxygen, sulfur, fluorine, and chlorine. These are in part elements whose principal valences vary greatly.

When simple negatively charged ions are linked to a central atom the most frequent coördination number met with is six; only four negatively divalent ions can generally be arranged around a central atom. Many researches have been conducted recently to study the factors which govern this number. The maximum coördination number in every case depends not only on the charge of the complex addendum, but also on the size of the central ion (59); when the same charges are concerned, the coördination number is smaller with a small central ion than with a larger one. Thus only four fluoride ions can be arranged around the small, trivalent positive boron ion ( $\text{KBF}_4$ ), while around the larger aluminum ion six can be arranged ( $\text{K}_3\text{AlF}_6$ ). The pentavalent nitrogen atom can add only three

oxygen ions, as in  $\text{K} \begin{array}{c} \text{O} \\ \diagup \quad \diagdown \\ \text{NO} \\ \diagdown \quad \diagup \\ \text{O} \end{array}$ ; phosphorus on the other hand can add four, as in  $\text{K}_3\text{PO}_4$  (77).

Also, the distance of approach to the same central ion will vary with the charge on this ion (47, 60, 85). In



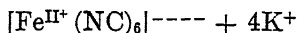
the distance of approach of the  $\text{NH}_3$  is 2.5 A.U., while in



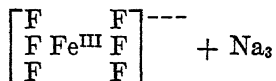
it is at a distance of 1.9 A.U. from the trivalent cobalt ion. This diminution of the distance indicates a very strong polarization; one might therefore assume that in this case the single electrons of the ammonia molecules are drawn so far toward the  $\text{Co}^{+++}$  ion that they belong equally to the coordinated constituent and to the central ion. Such compounds particularly concern us here, because they are either only very weakly paramagnetic or else diamagnetic.

*Dipoles or polarization in complexes*

In some complex ions, the coordinated ions are not uniformly surrounded by particles of opposite charge, but rather are subjected to a one-sided attraction to the central atom. We would therefore expect a one-sided polarization of this coordinated ion, which would give rise to dipoles, so that the bond between central ion and coordinated ion would be strengthened. For example, if one dissolves potassium ferrocyanide in water, the cyanide ions remain united with the central atom, while the potassium ions split off.



Judging from the magnetic data on this compound, the cyanide ions are linked to the central ion quite differently than, for example, are the fluoride ions in the compound



in which the iron ion shows the magneton number 29. The ferrocyanides themselves show no sign of the magnetism characteristic of the  $\text{Fe}^{++}$  (4.9 magnetons for the spin alone). A quite different type of linkage must therefore be present. From further experiments we shall see that one cyanide group of the hexacyanides can easily be attacked and replaced by neutral atoms, radicals, or molecules. If we had assumed an ionic linkage of the CN group as exists, for example, with the fluoride ions of  $\text{Na}_3\text{FeF}_6$ , this fact would be practically inexplicable. The assumption of a "true valence linkage" makes the matter less complicated. Unfortunately there is still lacking a definite explanation of the linkage in all of these complexes. In the case of potassium ferrocyanide we have clearly a case of "true valence linkage" giving rise to shared electron groups. In this substance we have the special case in which loss of the chemical properties



of the central ion occurs simultaneously with the loss of its paramagnetic quality. The same condition obtains in nickel dimethylglyoxime.

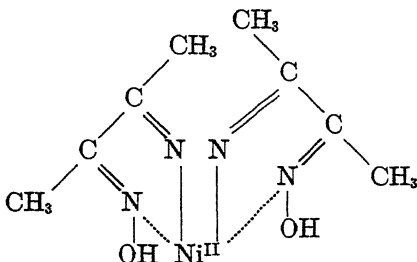
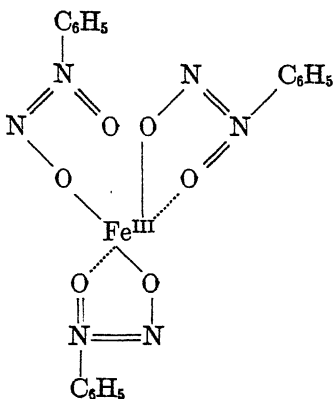
In complexes formed by addition of an uncharged dipole molecule to a central ion, this dipole molecule becomes polarized in one direction by the central ion, and the dipole moment is increased. Consequently the permanent electrical moment of this molecule is not alone sufficient to determine the strength of the linkage between a dipole molecule and a central ion. So, for example, the great affinity of pyridine for iron is not entirely due to the dipole nature of pyridine, but depends greatly upon the *specific nature of the pyridine nitrogen*. Again, the dipole moment of water is greater than that of ammonia. Were the binding energy determined only by the permanent dipole moment, then with any central ion the hydrate would always be more stable than the ammonia complex, and the latter with water would react to give ammonia and the hydrate. Experiment shows that though this often does occur, in many cases it does not.

As we have already mentioned, it is especially interesting that paramagnetic central ions of many complexes behave magnetically like simple ions, although chemically they are characterized as very strong complex compounds. So, for example, we have the iron salt of nitrosophenylhydroxylamine or iron benzohydroxamate and others, in which the iron ion is completely masked. On the other hand, there are other compounds in which the magneton number is low, thus in (33)

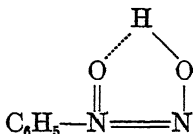


the Weiss magneton number is 4.3. The copper ammonia compounds show smaller susceptibility than the hydrates, a matter of interest, since because of the smaller dipole moment of the  $\text{NH}_3$  it was formerly expected to have less effect on the central ion. Then some salts such as  $\text{K}_4\text{Fe}(\text{NC})_6$ ,  $\text{K}_2\text{Ni}(\text{NC})_4$ , and  $\text{K}_2\text{PtCl}_4$  are diamagnetic, that is, the paramagnetic character of their metal ions has completely vanished.

From these few data, it may be seen that knowledge of the magnetism is of great importance in chemistry of the complexes, a fact strongly supported by recent work of Cambi and Szegő (38). Iron compounds can in general be divided into two groups: (1) ionized ferrous and ferric salts, and (2) complex iron compounds. The former are strongly paramagnetic with 27 and 29 Weiss magnetons, respectively. The complex iron compounds are weaker paramagnetically the more firm the complex linkage to iron is, i.e., the more they lose their residual affinity. In our investigation two compounds can be chosen as chemically and magnetically characteristic, viz., cupferron iron and nickel dimethylglyoxime.

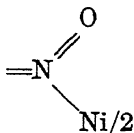


In both compounds the paramagnetic metal ion is linked to HNO groups as an inner complex salt. Both are strong complexes from a purely chemical point of view,—compounds in which the ionic nature of the metal, chemically speaking, is completely lost. By the study of their constitution, i.e., by exact knowledge about the linkage of the metals to the HNO group, we hope to be able to draw conclusions about their magnetic behavior. In cupferron iron, the metal ion is linked by principal valency to oxygen, since the hydrogen of the  $\text{=NOH}$  group is replaced. The iron is also joined by partial valences to oxygen since the pentavalent nitrogen



has no partial valences to be utilized. On the other hand, the doubly bound oxygen exerts a field force. The iron in this compound behaves magnetically like a ferric ion—its magneton number is 29.

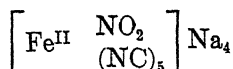
We find quite a different valence situation in nickel dimethylglyoxime. As Pfeiffer (66) has irrefutably shown, the nickel is linked with principal valences to the central ion



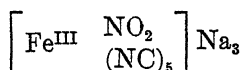
and with coördination valences to the  $\text{=N—OH}$  group. Since the  $\text{=N—OH}$  group contains trivalent nitrogen, this nitrogen atom still has partial valency to be utilized. The compound is diamagnetic and there-

fore, in the magnetic sense, a true complex salt. Both chemically and magnetically, nickel dimethylglyoxime, then, is a strong complex. In the undistorted complex the central ion can be recognized neither by ionic reactions nor by paramagnetism.

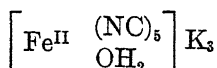
A thorough magnetic study of the pentacyano ferrous and ferric compounds led us to the observation, new in principle, that sodium nitroprusside (sodium pentacyano-nitroso-ferrate) possesses outstanding diamagnetic properties, while all other penta- and hexa-ferricyanides studied magnetometrically by Welo (81) were paramagnetic. Therefore, although it contains iron in the trivalent condition, it belongs magnetically to the complex ferrocyanides. In alkaline solution it is known that sodium ferropentacyanide-nitrite



is formed (49), which by oxidation by permanganate goes over to the ferric compound.



This in contrast to sodium nitroprusside is paramagnetic (14). Ferro-aquo salt

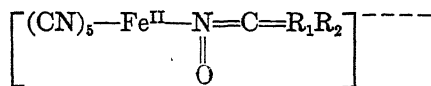


links nitric oxide, NO, similarly to carbon monoxide, CO, to the divalent iron through its coördination valences. A secondary oxidation results, however, and ordinary sodium nitroprusside with trivalent iron is formed.

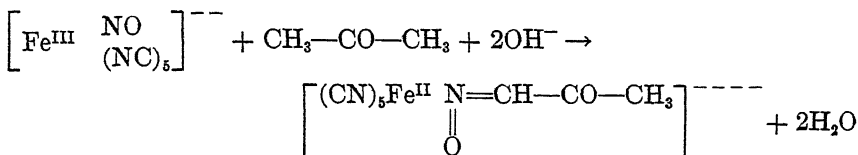
We assume that molecular oxygen, O<sub>2</sub>, also can be linked to ferro-aquo salt by chemical valence. The magnetic relations involved in molecular oxygen are much more complicated than they are in NO. Both, however, are paramagnetic gases. In the oxygen molecule two outer electrons behave extraordinarily like, although not to so great an extent as, the single electron of nitric oxide, which exists outside a system of ten electrons, and therefore behaves like a valence electron. We see, therefore, that the two gases bear a certain similarity in electron configuration as well as in magnetic behavior.

Sodium nitroprusside, because of the reactivity of its complex-bound NO group, is a valuable reagent in analytical chemistry. In particular, aldehydes and ketones can be recognized by vivid color reactions. Cambi (37) and collaborators have shown that reaction occurs with all substances having an activated —CH<sub>2</sub>— group, and have succeeded in isolating a

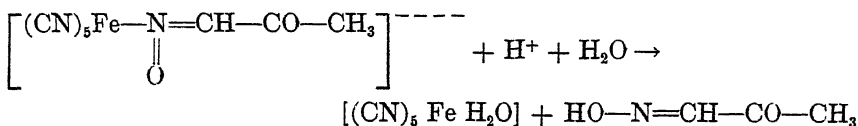
number of strongly colored compounds and explaining their structure. According to Cambi the organic compounds with sodium nitroprusside form deep red complexes; in the case of acetone,



formed by the reaction



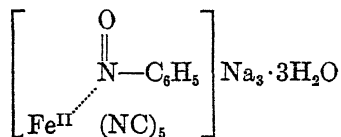
Thus we have occurring here a transformation quite similar to change from nitroprussi salt to nitroprusso salt in alkali; the deeply colored complexes, generally not very stable, are split by dilute acids to form isonitroso compounds and ferro-aquo salt:



which proves their structure.

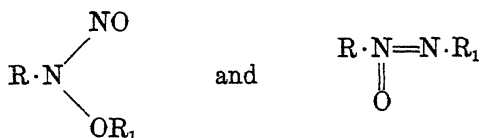
We may see from all of these examples what special interest sodium nitroprusside, and indeed all metal compounds of the iron family which link NO coördinatively, claim both practically and theoretically. It is very interesting further to note that NO can be displaced from most compounds by CO. We know that the iron in the blood pigment possesses this ability in the highest degree. In the metal carbonyls ( $\text{Fe}(\text{CO})_5$ ) and nitrosyls, NO and CO are joined to the metal atoms by coördination valences without regard to the valence ("electrovalence") of the metal.

When the principal valences are satisfied the radicals CO and NO are linked differently, as they are when a keto group ( $\text{R}=\text{CO}$ ) or a monovalent nitroso group ( $\text{R}=\text{NO}$ ) results from the saturation of principal valences, e.g., in nitrosobenzene. In spite of this, however, the characteristic coördinative linkage of the nitrogen to iron still operates, as we have proved by the synthesis (12) of the following compound:

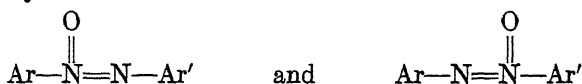


### III. DETERMINATION OF STRUCTURE OF NITROSOARYLHYDROXYLAMINES AND THEIR COMPLEX METAL SALTS

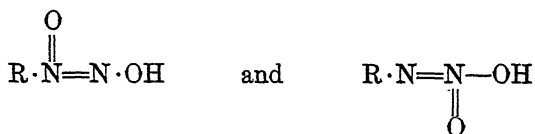
It has been found that nitrosoarylhydroxylamines can exist in two tautomeric forms. In some cases two isomeric esters have actually been isolated, for which the following structural formulas have been given.



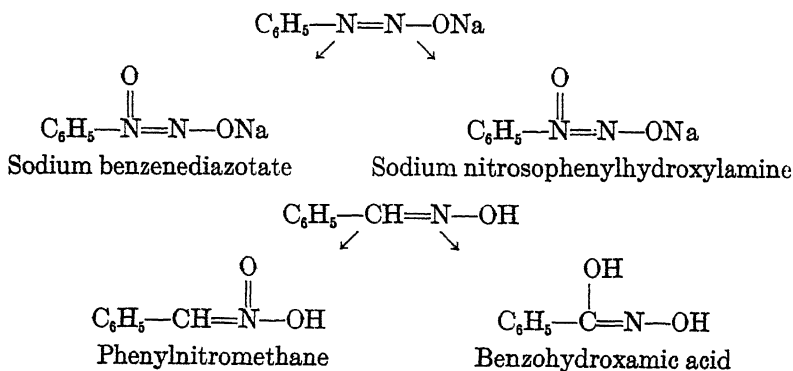
The work of Angeli (2) on the structure of the isomeric azoxy compounds forms the basis for the determination of the constitution of the nitrosoarylhydroxylamines. Angeli showed that azoxybenzene compounds of the general formula  $\text{Ar}-\text{N}_2\text{O}-\text{Ar}'$  exist in two isomeric forms which are represented by



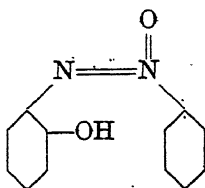
He formulated the diazo acids and the nitrosoarylhydroxylamines as follows:



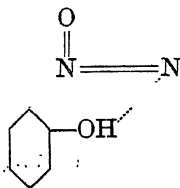
This structural representation was supported further by the work of Bamberger and Baudisch (6) on the oxidation of diazohydroxides and azo compounds. At the same time the analogy between oximes and diazo compounds, first recognized by Hantzsch, was broadened.







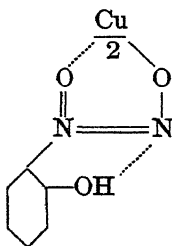
normal I



iso II

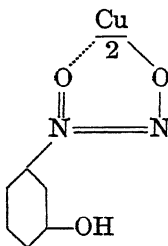
The "normal" *o*-hydroxyazoxybenzene, because of the free phenolic hydroxyl, is easily soluble in dilute alkali; the iso compound however is not, since the hydroxyl is anchored by subsidiary valences to the nitrogen. Compound I dyes wool and silk, but compound II does not.

If the above relations are applied to the nitrosoarylhydroxylamines we should expect the character of the nitrosohydroxylamine group to be modified when the linking ability of the trivalent nitrogen atom is changed by the influence of an *o*-hydroxyl group. This influence should be diminished in *m*- and *p*-hydroxy compounds, since these no longer offer the possibility for the formation of a six-membered ring. With this in mind we prepared the copper salts of the following three acids (24, 27, 28):



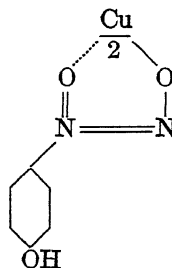
I

(clear gray)



II

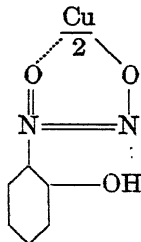
(greenish gray)



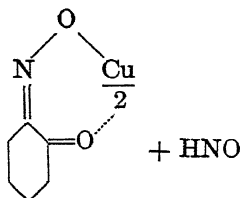
III

(grayish green)

The copper salt I is basically different chemically from II and III. Although II and III are stable salts insoluble in water, the clear gray copper salt I even on standing in the air splits  $\text{HNO}$  and is changed into the copper salt of *o*-quinone oxime, which is deep red in color (25).



I

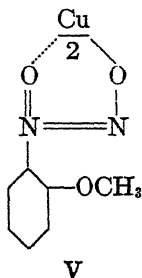


IV

+ HNO

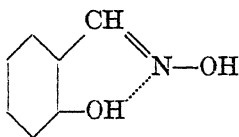
As the formula shows, the *o*-hydroxyl group in the salt I is coordinatively linked to the trivalent nitrogen atom. Because of this, the nature of the nitrosohydroxylamine group is completely changed, and there results the spontaneous decomposition into the quinone oxime copper salt and HNO. The formation of HNO is evidenced by the evolution of red-brown vapors of  $\text{N}_2\text{O}_3$ .

When the *o*-hydroxyl group is methylated, the corresponding copper salt (23, 29)

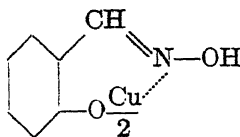


is a silver-gray metal compound, unchanged by boiling in benzene solution. Although the copper salt I is easily soluble in either water or organic solvents, the copper salt V is soluble only in benzene and a few other organic solvents.

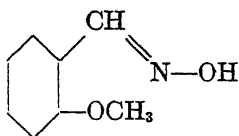
The differences in the solubility of the copper, cobalt, and nickel salts of the *o*-, *m*-, and *p*-hydroxynitrosophenylhydroxylamines (I, II, and III above) are particularly striking. While the ortho salts of the three metals are equally soluble in water and ordinary organic solvents, the meta salts are soluble only in acetone, and the para salts are insoluble in either water or organic solvents. Our assumption of the coördination linkage between the phenolic hydrogen and the trivalent nitrogen is also borne out by the work of Feigl (43) on the inner-complex formation by salicylaldehyde.



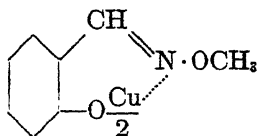
Free acid



Clear green copper salt



Forms no copper salt

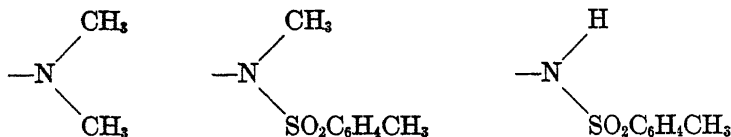


Dark brown copper salt

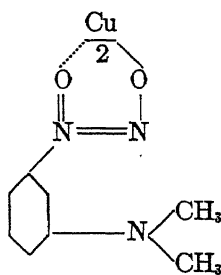


The analogy between the oximes and the nitrosohydroxylamines comes to light here.

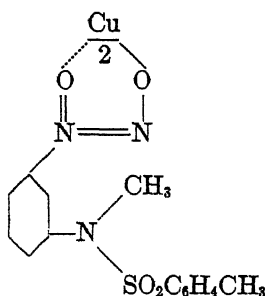
In order to obtain more exact insight into the character of the influence exerted by side chains, above exemplified by the *o*-, *m*-, and *p*-hydroxy-nitrosophenylhydroxylamines, we employed an amino group meta to the nitrosohydroxylamine group and varied only the properties of the nitrogen by substitution of the hydrogen by other residues:



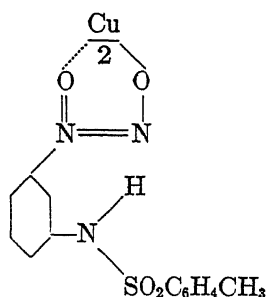
It was expected that the trivalent nitrogen atom in I would retain the coordination valence properties of the amino group,  $-\text{NH}_2$ , since in it the hydrogen atoms have been replaced only by methyl groups. However in II and III the group  $-\text{SO}_2-\text{C}_6\text{H}_5-\text{CH}_3$  is attached, which in itself possesses coordination valences, so that the coordination valence of the nitrogen should be partially neutralized and the characteristic properties of the nitrogen should disappear. These expectations are borne out by the properties of the following metal complex salts. The copper salts are again, in the case of these three acids, the most characteristic, therefore we can compare them to the copper salts of the hydroxy-substituted nitroso-phenylhydroxylamines.



I  
(yellowish  
green)



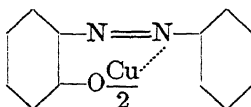
II  
(bright blue)



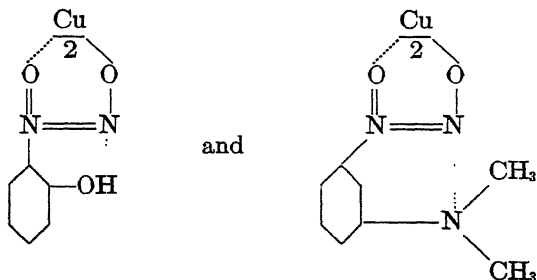
III  
(bright blue)

The influence of the nitrogen of the  $-\text{N}(\text{CH}_3)_2$  group of salt I on the trivalent diazo nitrogen is appreciably greater than in salts II and III, as shown by the green color of the salt I. For, according to Pfeiffer, hand in hand with the increase in the saturation of the subsidiary valences, there is observed a shifting of color from blue (color of the copper ions) through

violet and red to brown. The copper salt of *o*-hydroxyazobenzene, which is colored brown, is to be considered a strong complex (5).

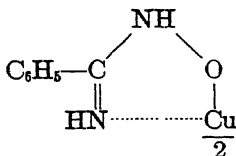


In the copper salt of *o*-hydroxyazobenzene the coördination linkage goes from copper to nitrogen, and affects the chemical and physical nature of the central copper ion. However, with the salts I, II, and III, we have made the new important observation that the inner-complex forming nitrosohydroxylamine group can be changed in its properties if the subsidiary valence of the side chain group is not linked to the central metal ion itself, but to the diazo nitrogen atom instead. We have already assumed a similar constitution for the *o*-hydroxynitrosophenylhydroxylamine salt:

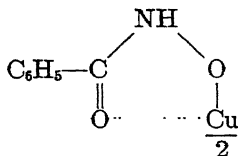


Not only the colors of the copper salts I, II, and III, but also the colors of the iron salts differ. While the iron salts of the acids from II and III are red-brown, and outwardly resemble cupferron iron salt, the iron salt of *m*-dimethylaminonitrosophenylhydroxylamine, I, is brownish black (with violet luster). The copper salt II is easily soluble in hot chloroform and benzene, difficultly soluble in hot alcohol, and insoluble in acetone, ether, and water. It dissolves in acids with a dark red color.

A few additional instructive examples will be given showing the effect of coördinative linking with respect to copper salts of nitrogen and oxygen compounds.



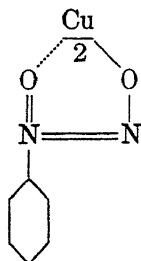
Benzamide oxime copper salt  
(brown)



Copper benzohydroxamate  
("grass" green,  
bright green)

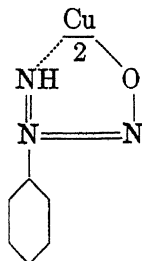
The nature of the complex depends upon whether the subsidiary valences of the central metal ion are linked to nitrogen or oxygen atoms; judging by the color, the nitrogen complexes are much stronger. Magnetic measurements of these salts would be of special interest, particularly of the iron, cobalt, and nickel salts.

Another very interesting illustration is the comparison of the copper salt of nitrosophenylhydroxylamine with that of nitrosophenylhydrazine:



I

Nitrosophenylhydroxylamine  
copper salt  
(grayish blue)

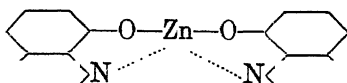


II

Nitrosophenylhydrazine copper  
salt  
(deep red-brown)

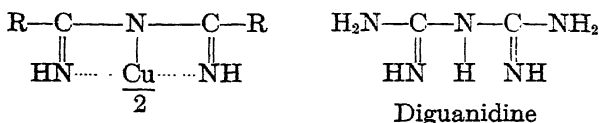
In all cases where the coordinative linkage is strengthened, the color of the copper salt is shifted toward red and brown. The basis for representing the salts I and II with analogous structures is the fact that salt II, by simple treatment with a mixture of acetic acid and alcohol, is converted into the grayish blue salt I, i.e., the imino group is replaced by an oxygen atom. While the salt I is relatively stable, the nitrosophenylhydrazine copper salt decomposes explosively, often without cause, with deposition of metallic copper.

The inner-complex salts of 8-hydroxyquinoline are also interesting. For example, the zinc salt has the following structure (41):

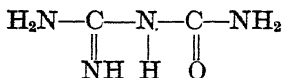


Coördination valences link the cyclic nitrogen of quinoline to zinc, while the metal is linked by principal valences to oxygen.

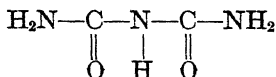
Other interesting copper salts in which the metal is linked to nitrogen by principal valences are, according to Ley and Müller (57), those of the amidines, which possess the following structure:



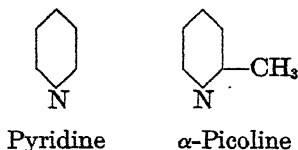
If we replace the  $\text{>C=NH}$  group of diguanidine by carbonyl, we produce dicyandiamidine,



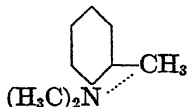
and the important biuret (76):



In concluding the discussion of coördination involving the nitrogen atom and of ways of affecting the reactivity of nitrogen complexes, we should like to mention a few additional examples which illustrate chemical changes of the complex-forming group. We have found that the coördination valences of the nitrogen in pyridine, which show great affinity for divalent iron, are very much weakened by the presence of an *o*-methyl group. We have explained this as a diversion or neutralization of the coördinative valence (13):



Quite similarly we explain the unreactivity of the *o*-methyl tertiary amines (51):



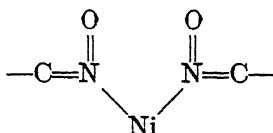
The nitrosophenylhydroxylamine ammonium salt, which was introduced to analytical chemistry under the name "cupferron" by Baudisch (9, 26), has served for many uses in the quantitative separation of metals, as the great number of publications testify. In neutral solution it precipitates



although chemically the properties of the central iron ion have completely disappeared, and in acetone solution it does not conduct the electric current. There still remain unexplained relations between the complex addenda and the central metal ion which influence the electron configuration and especially the spin of the electrons.

A series of nickel salts with dimethylglyoxime are known which have red, blue, green, and yellow colors. We shall discuss their constitution here.

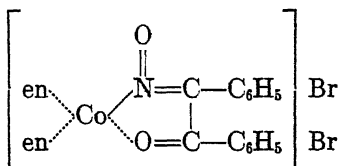
The color-determining portion of the molecule in the nickel glyoximes is the atom grouping



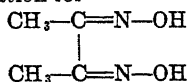
in which it makes no difference whether the two  $\begin{array}{c} \text{O} \\ \parallel \\ -\text{C}=\text{N}- \end{array}$  groups belong

to the same dioxime, or to two different dioximes. The atom configurations of the differently colored nickel dimethylglyoxime compounds have been cleared up by the work of Meisenheimer, Pfeiffer, and Thilo (75).

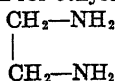
T. Meisenheimer has shown that the  $\alpha$ -dioximes, formerly expressed in the *syn* form, possess the *anti* configuration. Therefore on the basis of purely steric relationships, it must be incorrect to assume that the nickel atom is generally linked to *oxygen*, if we maintain that the  $\text{DH}_2$ <sup>4</sup> possesses in complex salts two places of coördinative linkage. The important experimental proof of the duplexity of the dioxime in inner-complex compounds was furnished by P. Pfeiffer and I. Richarz. They prepared<sup>5</sup>



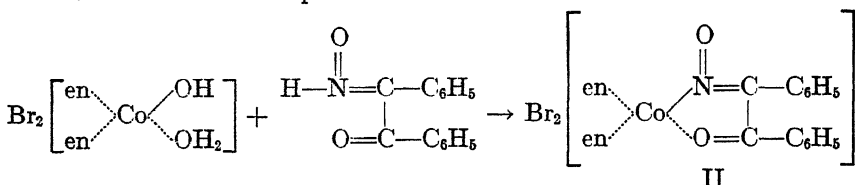
<sup>4</sup>  $\text{DH}_2$  is used as the abbreviation for



<sup>5</sup> "en" is used as the abbreviation for ethylenediamine:

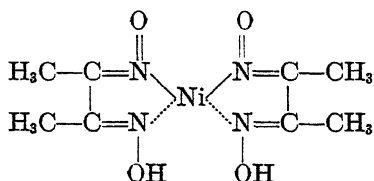


an orange-yellow compound from  $\alpha$ -benzil monoxime in the following way:  $\alpha$ -benzil monoxime was reacted with *cis*-hydroxo-aquo-diethylenediamine with formation of the compound II:

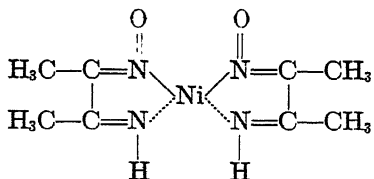


However, the  $\beta$ -benzil monoxime under exactly the same conditions gave no such salt. In salt II the monovalent rest  $\text{R}-\text{CO}-\text{C}(\text{R})\text{:}\dot{\text{N}}\text{:O}$  quite certainly contains two positions of coordinative linkage: only so can the coordination number six of cobalt be fulfilled. II is therefore a true inner-complex salt.

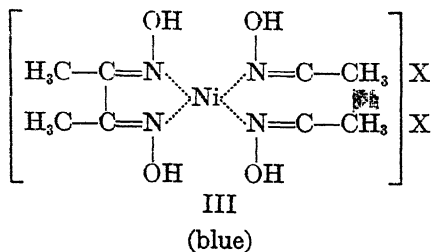
On the basis of these investigations, Pfeiffer formulated the red nickel glyoxime as the inner complex salt:

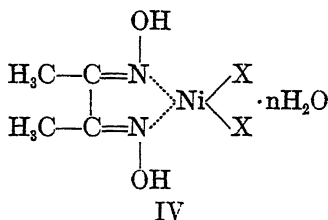


The final proof of this concept was given by P. Pfeiffer by preparation of the difficultly water-soluble nickel diacetyl iminoxime, which crystallizes in red plates:



By addition of acids to the  $\text{Ni}(\text{DH})_2$ , blue and green salts are obtained, which are formulated as follows:

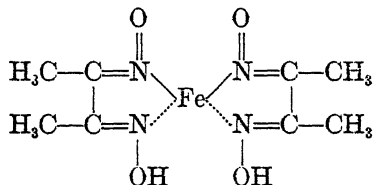




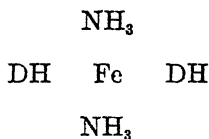
(bright green)

In III, by addition of acids to the  $\text{Ni}(\text{DH})_2$ , the two atom complexes have been transformed back to the original oxime groups. In IV the compound contains a strongly acid residue.

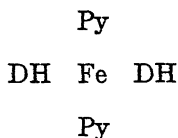
Dimethylglyoxime also gives complex salts with iron. Thus, for example, a neutral ferrous salt solution containing dimethylglyoxime is colored deep red upon addition of ammonia. The solutions are very unstable and decompose on standing in the air, with precipitation of ferric oxide hydrate. From the color, which is very similar to that of the nickel dimethylglyoxime, one might suppose that there is formed in the solution the compound:



This may be abbreviated as  $\text{DH}-\text{Fe}-\text{DH}$ . However, the constitution is probably:



for, as L. Tschugaeff showed, if one treats ferrous salt solutions containing dimethylglyoxime with an excess of pyridine instead of ammonia, after a short time there is precipitated a compound



in the form of dark brownish red crystals.



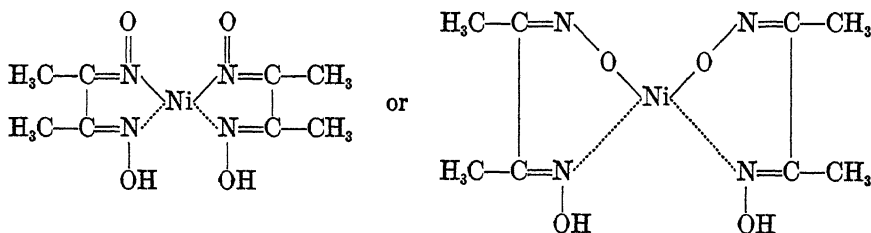
Little is yet known about the compounds of trivalent iron with dimethylglyoxime. From both divalent and trivalent cobalt a great series of characteristic complex compounds with dimethylglyoxime are known.

From the standpoint of the metal atom (iron, cobalt, nickel) it is found that iron occurs only as a hexavalent coördinative central atom. With cobalt the compounds with coördination number six are far more stable; nickel (also copper) gives no compounds containing a hexavalent metal atom, and only those with four coördination linkages are known.

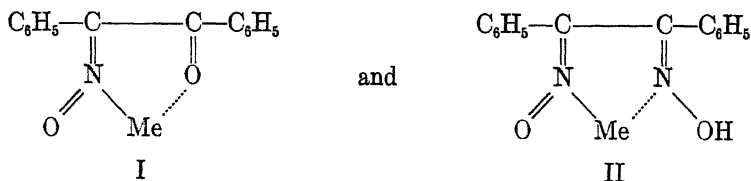
In general, it appears that the chemical behavior of the metals iron, cobalt, nickel, and copper in their complex dioxime compounds is completely similar. All replacements are continuous, yet clearly graduated however. Thilo finds no indication of any specific affinity of dimethylglyoxime for nickel.

Later we shall go more into detail about the magnetic properties of nickel dimethylglyoxime.

Next we are concerned with the chemical configuration and its relation to the linkage of the central atom. It is of fundamental importance, in determining the properties of the metal complex compound, whether the metal is linked directly to nitrogen or to oxygen:



The assumption of the "nitron formula" is necessary to explain the fact that only the  $\alpha$ -form (I) of the benzil monoximes and only the  $\alpha$ -stereoisomer (II) of the benzil dioximes leads to the formation of inner complexes (68).



From the investigations on the structure of nickel dimethylglyoxime reported here, it is seen that in the *red* nickel dimethylglyoxime the central ion is linked by principal valency to the complex addendum. In the *blue* salt on the other hand, the =NOH group is linked exclusively by sub-

sidary valences to the nickel ion. Iron behaves in a quite similar manner in the dimethylglyoxime compounds.

V. THE RELATION BETWEEN THE STRUCTURE OF CERTAIN WERNER IRON COMPLEXES AND THE IRON OXIDE HYDRATES AND IRON OXIDES FORMED FROM THEM

It has been demonstrated by the experiments of Welo and Baudisch (83) that under certain conditions in the oxidation of magnetite ( $\text{FeO} \cdot \text{Fe}_2\text{O}_3$ ), the inner lattice structure of the ferromagnetic magnetite is retained and a ferric oxide ( $\text{Fe}_2\text{O}_3$ ) is formed which is ferromagnetic. It is a surprising fact that magnetite after complete oxidation to iron oxide ( $\text{Fe}_2\text{O}_3$ ) retains its spinel structure.

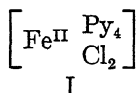
They found further that ferrous hydroxide,  $\text{Fe}(\text{OH})_2$ , in a strongly alkaline solution quickly changes by loss of water into ferrous oxide, which through oxidation by air is converted chiefly into the rhombohedral iron oxide, hematite ( $\alpha\text{-Fe}_2\text{O}_3$ ), along with small amounts of  $\gamma\text{-Fe}_2\text{O}_3$  (31). In this case, then, during the oxidation, intermediate formation of magnetite does not occur, or occurs only to a slight extent as disclosed by the production of small quantities of ferromagnetic iron oxide with the larger quantities of  $\alpha$ -oxide. The nature of the process of formation of  $\alpha$ - and  $\gamma\text{-Fe}_2\text{O}_3$  was thereby explained by us through x-ray determination of the crystal structure. Always in those cases in the oxidation of ferrous hydroxide or oxide in which magnetite is formed as an intermediate, complete oxidation leads to the formation of a ferromagnetic iron oxide. On the other hand, direct oxidation of ferrous oxide to ferric oxide leads to the formation of non-magnetic hematite.

However, there appeared to be an exception to this rule, since, as Albrecht (1) first showed, it is possible by simple dehydration of artificially prepared  $\gamma$ -ferric oxide hydrate ( $\gamma\text{-Fe}_2\text{O}_3 \cdot \text{H}_2\text{O}$ ) to prepare a ferromagnetic  $\gamma\text{-Fe}_2\text{O}_3$  which has a crystal structure identical with that of  $\gamma\text{-Fe}_2\text{O}_3$  obtained from magnetite. Here, the intermediate formation of a ferrous-ferric oxide cannot be postulated, since with the  $\gamma$ -hydrate we have from the beginning trivalent iron.

The preparation of  $\gamma$ -iron oxide hydrate in very pure form by Baudisch (17) and by Baudisch and Albrecht (20) by a new method, afforded the opportunity to investigate more closely the inner lattice structure and other properties of this interesting compound. Two principal factors seemed to favor the production of  $\gamma$ -hydrate by the oxidation of solutions of divalent iron salts; firstly, the pH, and secondly, the tendency of the reacting system to form more or less stable Werner complexes with iron as the central atom. The pH may be of importance only inasmuch as it affects the dissociation equilibria of the intermediate labile complexes. It was our desire to

demonstrate the existence of such labile complex compounds, which might contain both divalent and trivalent iron. This expectation might be difficult to realize, since such complexes would probably in general exist in solution only as components of a dissociation equilibrium and not be readily obtained for analysis. Our experience with the preparation of  $\gamma\text{-Fe}_2\text{O}_3$  pointed to the supposition that the  $\gamma$ -hydrate is formed only when intermediate compounds are formed which have a constitution analogous to that of magnetite. Magnetite itself is ferromagnetic also in the *hydrated* condition and thereby fundamentally different from the  $\gamma$ -ferric oxide hydrate. Further x-ray investigations showed that hydrated magnetite gives the x-ray spectrum of the water-free magnetite ( $\text{FeO}\cdot\text{Fe}_2\text{O}_3$ ), and therefore is entirely different from the Roentgen spectrum of the  $\gamma$ -ferric oxide hydrate which was first determined by Böhm (36).

Since the  $\gamma$ -iron oxide hydrate prepared by oxidation of ferrous salts is a true hydrate, the hypothetical intermediate compounds with divalent and trivalent iron could not be hydrated magnetite. We obtained evidence as to the mechanism of the formation of the  $\gamma$ -hydrate only when we studied its synthesis from a simply constituted Werner complex salt. We chose for this study the canary-yellow complex salt easily prepared from ferrous chloride and pyridine, having the following constitution:



This complex salt was first described by N. Costachescu and G. Spacu (39). Just recently in an extensive compilation of data which relate partly to the dissolved state (conductivity studies), and partly to the crystalline state (x-ray and magnetic studies), it has been shown that complex salts of halides of the transition elements with four neutral amine molecules, contain the halogen in general still in the inner sphere.

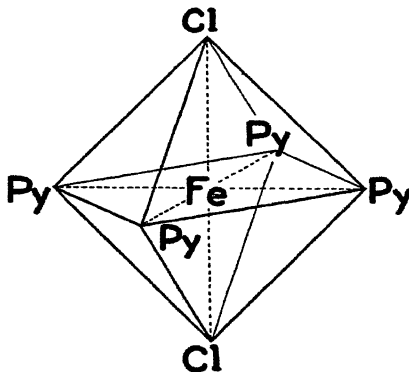
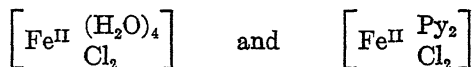
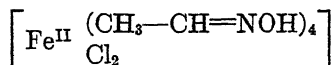


FIG. 1

To these "pseudo salts" (as they are termed by Hantzsch) the compounds

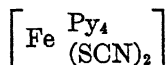


also belong, according to the recent investigations of Stelling (74). The compounds of ferrous chloride with acetaldoxime and other aldoximes also belong to this group (48); thus, for example,



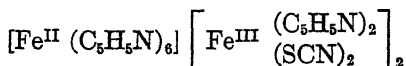
This fact is of interest in connection with our concept of the genetic relationship of hydroxamic acid and aldoximes to pyrrole and pyridine compounds mentioned earlier.

Of the tetraamino complexes mentioned, the thiocyanate compounds especially have been thoroughly investigated recently. The tetrapyridine compound of ferrous thiocyanate,

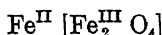


II

which was first prepared by Grossmann and Hünseler (45), is autoxidizable in solution. A. Rosenheim, E. Roehrich, and L. Trewendt (71) have recently studied the autoxidation of solutions of this compound and found that a bimolecular compound is produced which contains both divalent and trivalent iron, and is constituted analogously to magnetite:

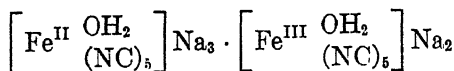


Spacu (73), who repeated the work of Rosenheim and collaborators, obtained in his many experiments analytical values which varied quite widely, yet which were in part in agreement with those of Rosenheim. He expresses the view that in all of these intermediate products from the autoxidation of II, mixtures of tripyridine ferric thiocyanate and closely related ferrous compounds are present. Rosenheim (70) also concurred in this view. At any rate, it is certain that the yellow compound II is autoxidizable and compounds can be formed which are constituted analogously to magnetite.



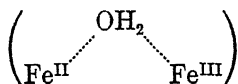
In direct connection with these results stands our earlier work with the pentacyano-aquo compounds. The yellow pentacyano-aquo-ferroate is

autoxidizable in solution and forms polynuclear complexes which contain both divalent and trivalent iron (40). According to Michaelis (61), a pH of 6.6 furnishes the best condition for the formation of a salt having the structure



### III

If a yellow solution of ferro-aquo salt is mixed with a violet solution of ferri-aquo salt in the molecular ratio 1:1, the new salt, III, is formed which is brownish yellow. This salt is also formed when potassium ferrocyanide solutions are exposed to light in the presence of oxygen. As we have found, the formation of the bimolecular salt is inhibited when small amounts of sodium azide are added to the ferro-aquo salt solution. This compound possesses an affinity only for the coördination position occupied by the water molecules, and prevents the water molecule from forming a bridge

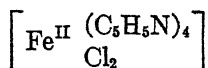


and thus joining two molecules so that they can form a bimolecular complex. Therefore, when a yellow ferro-aquo salt solution is oxidized (by oxygen or hydrogen peroxide) in the presence of sodium azide, a deep violet-colored solution is obtained, which contains the ferri-aquo salt in equilibrium with the corresponding ferri-azido salt (15).

If we apply these principles to the autoxidation and hydrolytic decomposition of tetrapyridine ferrous chloride, we reach some useful conclusions as to the mechanism of the formation of  $\gamma$ -oxide hydrate. These will be discussed below.

We have already shown in earlier publications (20) that from ferrous chloride solutions treated with pyridine, by passage of air or oxygen a certain amount of orange-yellow  $\gamma\text{-Fe}_2\text{O}_3 \cdot \text{H}_2\text{O}$  is precipitated. On the other hand, if ferric chloride is treated under the same conditions with pyridine, a red hydrate is obtained, which on dehydration gives non-magnetic hematite. By varying the conditions of temperature and precipitation, amorphous orange-yellow hydrates are obtained, which on dehydration also give only hematite.

In further experiments we used the yellow pyridine salt,



which dissolves to a colorless solution in water. These solutions in the presence of air are extremely unstable and hydrolyze with the splitting-out of pyridine. The colorless solutions become green and turbid, and orange-yellow ferric oxide hydrate is quantitatively precipitated. The filtered clear solution shows the pH 5.6–5.8. If the yellow salt is dissolved in a strong aqueous solution of pyridine, the primary green coloration does not appear, but the solution which is at first colorless, becomes dark brown and finally, after long shaking with air, all the iron is precipitated as an orange-yellow hydrate. The pH of the clear colorless filtrate is 7.1–7.3.

From these experiments it is seen that in both weakly acid and weakly basic solutions, autoxidation produces an orange-yellow hydrate. The dehydration of this hydrate at 260°C. yields a strongly ferromagnetic  $\gamma$ -Fe<sub>2</sub>O<sub>3</sub>. If, however, the yellow salt is dissolved, instead of in water, in superoxol (H<sub>2</sub>O<sub>2</sub>), vigorous oxidation ensues and heat is evolved. By dilution with water, a brown hydrate is obtained which is completely non-magnetic. If the hydrate is dehydrated at 260°C., a red oxide (hematite) is formed which is also non-magnetic. Thus, the result of the oxidation of the yellow salt is absolutely different, depending upon whether the oxidation is carried out with hydrogen peroxide or with air.

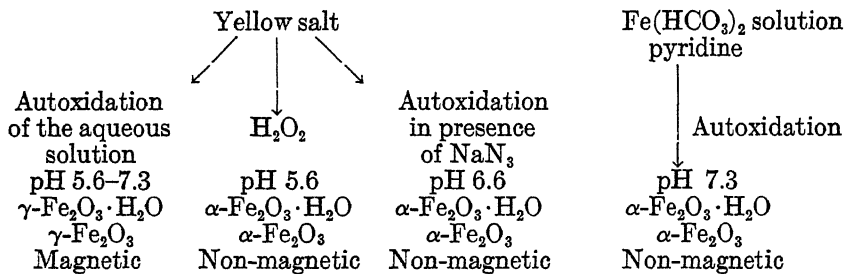
In order to effect oxidation as rapidly as possible with oxygen the salt was added in small portions to water saturated with oxygen, through which oxygen was continuously bubbled. There was formed almost immediately an orange-yellow hydrate. The dehydration of this hydrate also gave an  $\alpha$ -oxide (hematite) which contained only small amounts of admixed ferromagnetic oxide. The experiments were often repeated and showed conclusively that slow autoxidation of the yellow salt leads to the formation of the  $\gamma$ -hydrate, while rapid oxidation favors the formation of the  $\alpha$ -hydrate.

From the previously mentioned facts, it seems logical to assume that by autoxidation (slow oxidation) bimolecular complexes containing divalent and trivalent iron are formed which possess a structure analogous to that of magnetite, and by hydrolysis and oxidation are converted to the  $\gamma$ -hydrate. With *rapid* oxidation of the yellow salt, such ferrous-ferric complexes cannot be formed, and therefore only the  $\alpha$ -hydrate is obtained. Likewise by hydrolysis of the complex pyridine ferric compounds only  $\alpha$ -hydrate can be produced.

From some previously mentioned experiments with the pentacyano iron compounds we know that we can inhibit the formation of bimolecular complexes by sodium azide. We can therefore test our assumption of the intermediate formation of such complexes in the case of the tetrapyridine ferrous chloride compound. We would expect that in the presence of sodium azide only an  $\alpha$ -hydrate would be obtained by autoxidation of the

yellow salt, since sodium azide prevents the formation of polynuclear compounds and it is known that direct oxidation of divalent iron leads to the formation of the  $\alpha$ -hydrate. This is actually what does occur. If the yellow salt is dissolved in a dilute sodium azide solution, the autoxidation occurs in the same way that it does in pure water. The orange-yellow hydrate formed is not a  $\gamma$ -hydrate, however, but is a pure  $\alpha$ -hydrate. (The pH of the clear filtrate is 6.6.) Although the oxidation is quite gradual, and all other conditions are as given for the formation of the  $\gamma$ -hydrate, only the  $\alpha$ -hydrate is formed. We are thus able to prepare either  $\alpha$ - or  $\gamma$ -iron oxide hydrate from the yellow salt, depending upon the conditions of precipitation.

If instead of the yellow salt, a solution of ferrous bicarbonate is used, and it is oxidized slowly with oxygen in the presence of pyridine (pH = 7.3), in all cases  $\alpha$ -hydrate alone is formed. The ferrous bicarbonate has no tendency to form complexes with pyridine, therefore oxidation always converts the ferrous ions directly into the  $\alpha$ -hydrate.



It is hardly to be doubted that under the experimental conditions which lead to the quantitative formation of an intermediate bimolecular compound of magnetite structure, a quantitative yield of  $\gamma$ -Fe<sub>2</sub>O<sub>3</sub>·H<sub>2</sub>O is obtained. In case the ratio of the complex pyridine ferric chlorides to the pyridine ferrous chlorides in the mixture is different, then mixtures of  $\alpha$ - and  $\gamma$ -hydrates will be obtained. By quantitative determination of the ferromagnetism of the oxides formed by dehydration of the hydrates produced from the yellow salt, it should be possible to control the autoxidation process and the composition of the intermediate products. Formerly it was only possible to do this by tedious determination of the atomic ratios of the iron, nitrogen, and halogen.

That under the conditions mentioned (pH from 5.6-7.3)  $\alpha$ - and  $\gamma$ -hydrates and not the oxides, are formed, probably is to be traced to the fact that ferrous hydroxide, Fe(OH)<sub>2</sub> or FeO·H<sub>2</sub>O, remains as such in the intermediate ferrous-ferric compounds. In the course of the oxidation, the ferrous iron of the intermediate bimolecular complex is oxidized without

loss of the chemically combined water of hydration. Therefore we obtain by autoxidation at a pH slightly on the acid side or near the neutral point on the basic side, *hydrates*, not oxides.

If the aforementioned experiments with the yellow salt are carried out in exactly the same manner in a strongly alkaline medium (e.g., with piperidine) we no longer obtain hydrates, but rather oxides, or mixtures of the two. For example, if the yellow salt is put in piperidine and then diluted with water and shaken with air, there is intermediate formation of black magnetite, which on further treatment with air is completely oxidized to  $\gamma$ -iron oxide. The  $\gamma$ - $\text{Fe}_2\text{O}_3$  formed possesses a beautiful velvety red color. On short heating to  $260^\circ\text{C}$ . the oxide loses its magnetism, and the color deepens. The intermediate formation here of finely divided black magnetite is in agreement with the supposed formation of the complex intermediates analogous to magnetite when the autoxidation of the yellow salt occurs in weakly alkaline solutions.

The oxidation of the yellow salt with hydrogen peroxide in a strongly alkaline solution does not lead to the formation of the  $\alpha$ -hydrate, but instead to mixtures of considerable amounts of  $\alpha$ -oxide with small amounts of  $\gamma$ -oxide. Likewise, as we have already mentioned, the oxidation of aged  $\text{Fe}(\text{OH})_2$ , in other words, of  $\text{FeO}$ , leads to the formation of  $\alpha$ - $\text{Fe}_2\text{O}_3$ , as we have proved by x-ray analysis.

From these experiments we may therefore draw the conclusion that the process of formation of the  $\gamma$ -hydrate is not fundamentally different from that of the formation of the  $\gamma$ -oxide.

When we work with Werner complexes we have from the start a spatial orientation, whether the solid salts or their solutions are used. We have here an interesting case in which the spatial arrangement and the chemical composition of the intermediate complexes determine the type of lattice from which the hydrates and oxides that are formed are built up.

Besides differences in the lattice types of the iron oxides, we have to deal with differences in the surface properties, which are primarily governed by the "sensitive crystal properties" (empfindlich Krystalleigenschaften). In the following section we shall be principally concerned with the surface properties of iron oxides in their application to biochemistry. In order to determine by physical measurement the sensitive crystal properties, we have applied the measurement of the hysteresis and coercive force of the magnetic oxides, as we shall report in the second part of this review. We are therefore able to determine the fine structures of our iron oxides by x-ray measurement, and through the application of magnetic measurements, their hyperfine structures.

Since we may look upon the Werner complexes as a stage preceding the formation of submicroscopic crystals or colloidal particles, there is a logical



reason for treating the biological experiments with both types of iron compounds simultaneously in this review.

#### VI. THE IMPORTANCE OF ACTIVE IRON (WERNER COMPLEXES AND IRON OXIDES OR HYDRATES) IN BIOLOGY

The reactions at the surfaces of solid substances must be similar to the reactions with similarly constructed Werner complex salts. On the surfaces of solid iron oxides there exist coordination centers in the form of loosely bound iron atoms which are able to draw in radicals or even whole molecules to form primary coordinative bonds. Then the exchange or displacement occurs by which the reacting substances become activated. In the active states, reactions which would otherwise proceed extremely slowly may be catalytically accelerated or may even first become at all possible.

It is therefore not only the number of atoms on the surface that determines the activity but there must exist, as we have indicated, a special property of individual active atoms which affects the adsorption and, because of that, the activity. The proportion of these atoms in a special condition to the total number present is to be regarded as the measure of activity.

The catalytic activity of iron oxide varies greatly with the method of preparation, and our first problem is to prepare it of the highest possible activity. Then we may proceed with the biological tests. Researches of this type are still in a preliminary stage and there is, no doubt, room for much development and improvement.

The first direct experiments to test the biological activity of iron oxides and complex iron compounds were those of Webster (78, 79). He established the fact that red corpuscles, which are necessary for the growth of *Bacterium lepi-septicum* of virulent type D, may be replaced by sodium penta-cyano-aquo-ferroate ("aquo salt"). This typical Werner complex salt functions even when diluted to 0.002 mg. per cubic centimeter. Since these very dilute solutions gave benzidine reactions (blood test), the biological activities of both the iron compounds and the red corpuscles were measured as to their peroxidase properties, using the benzidine reaction as an indicator. As a result of these experiments of Webster, to whom we had supplied the so-called aquo salt, it appeared that *B. lepi-septicum* is to be regarded as an organism so delicately adjusted to oxygen pressure that it can grow prolifically and possess the specific properties of the virulent type only when oxygen is mechanically removed, or when the effects of oxygen are suppressed by the addition of catalase (blood, aquo salt).

This important observation of Webster concerning the biological action of aquo salt led us to examine our iron oxides for peroxidase properties by

means of the benzidine test. Since the oxides gave the typical blue coloration in a very distinctive manner we assumed, at first, that their biological activities could be characterized in a simple way by the color reaction. Further investigation showed that the relations are not so simple. Unfortunately, however, our first statements regarding the benzidine activity of iron oxides, which were made without a knowledge of the physical and chemical relations to the fine structure, have been expanded far beyond legitimate limits by other investigators, so that much confusion has arisen regarding the meaning of "active iron" (18).

Since the term "active iron oxide" has been applied also, in inorganic chemistry, to oxides having free energy and which approach a stable condition on aging, it is necessary to emphasize that when we speak of "active iron" we have in mind only those oxides or those complex iron compounds which have a favorable influence on the length of life, the growth, the virulence, or, generally, on the biological activity of bacteria and tissue cells.

Although we had previously carried the knowledge gained in our experiments on nitrate and nitrite reduction over into the biological field, and had shown that there is a connection between the iron content of highly virulent cholera bacteria and the power to reduce nitrates, our systematic experiments to obtain a deeper understanding of active iron compounds were first begun after Webster's positive results with aquo salt became known.

In aquo salt we had at hand a biological catalyst of known constitution, and we were justified in the hope that we could determine the nature of the so-called X mechanism of hemophilic bacteria. Avery (3) and his co-workers had found that a so-called X factor is necessary for good growth of hemophilic anaerobic bacteria and of pneumonia bacilli at atmospheric pressure. They found that red corpuscles do not lose their X factor quality even after heating in steam at 120°C., and they were finally able to replace the X factor by adding gum arabic-ferrous sulfate solution to the culture medium.

The extensive experiments of Webster and Baudisch (79) showed that aquo salt diluted to 0.01 mg. per cubic centimeter can replace the X factor with hemophilic bacteria. Influenza bacilli also grew in the presence of aquo salt, as did *B. welchii* and *B. sporogenes*. The experiments extended over a period of a year, and could always be repeated with the same results.

It was but a short step to go from sodium penta-cyano-aquo-ferroate (aquo salt) to the active iron oxide,  $\gamma\text{-Fe}_2\text{O}_3$ , since both compounds gave positive benzidine reactions, showed definite catalase properties, and absorbed oxygen in aqueous solutions or suspensions. If  $\gamma\text{-Fe}_2\text{O}_3$  were heated to 550°C. it changed its structure and lost these properties almost completely. No observable chemical change occurred when the oxide was

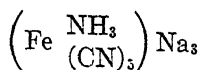
heated, for there was no loss of weight and chemical analysis again gave  $\text{Fe}_2\text{O}_3$  (inactive  $\alpha\text{-Fe}_2\text{O}_3$ ). The biological tests proved to be quite striking. While the  $\gamma\text{-Fe}_2\text{O}_3$ , similarly to the aquo salt, functioned as a very active X factor (biocatalyst, growth catalyst), the results with  $\alpha\text{-Fe}_2\text{O}_3$  were entirely negative.

It can be readily understood that we, at that time, attributed the surprising results primarily to the differences in crystal structure. We were then at the threshold of our oxide studies and as yet knew nothing of their great variability. Also, there was little that was definite about them to be found in the literature. This much, however, was definitely established by our experiments: the requirements of the X mechanism can be satisfied not only with red corpuscles but also with those iron compounds which show peroxidase properties and which absorb oxygen.

Our description of this biocatalytic action of certain iron oxides and of aquo salt excited the interest of other investigators and our experiments were variously repeated with both positive and negative results. When Miss Morris and Baudisch set about to repeat the Webster-Baudisch experiment half a year later, at the Rockefeller Institute, they at once met with difficulties, although both the media and the cultures were obtained from the laboratory of Webster.

In order to account for the discrepancies it seemed advisable to make a closer study of the aquo salt and its relation to the culture medium. The principal reason was Webster's proof that the ammino and nitro salts were far less active than aquo salt. In the ammino and nitro salts,  $\text{NH}_3$  and  $\text{NO}_2$  groups, respectively, replace the  $\text{H}_2\text{O}$  group in aquo salt. Dubos soon found a biological reaction with which to test the activity of aquo salt in the medium. This method could also be applied to iron oxide and proved to be far superior in sensitivity to the chemical tests for catalase.

The Dubos (15, 16, 22) test involves the longevity of pneumonia bacilli. Pneumococci die very rapidly in ordinary bouillon, but the addition of a small amount of blood will greatly extend their lives. There is, here, a certain analogy to the X mechanism. With pneumococci growing in flat dishes the presence of peroxides can be directly demonstrated with the benzidine reaction (benzidine + hydrogen peroxide). The peroxides shorten the lives of the cells and even kill them outright. The addition of catalase serves to keep them alive. We have here, then, a possible method of testing the catalase activity of aquo salt in bouillon and of comparing its activity with that of the ammino salt.



The method is as follows: 5 cc. of ordinary bouillon in a test tube is inoculated with a 12-hour old avirulent pneumococcus culture, strain RN derived from Type II, and the sterile salts are added. The tubes are kept at 37°C. for 24 hours and are then left to stand in the dark at room temperatures. From time to time, samples are removed and spread on blood agar.

The result was that aquo salt decidedly lengthened the lives of the pneumococci even in dilution to  $10^{-3}$ , whereas there was no effect with ammino salt. This experimental procedure was applied to test the catalase action of iron oxides. Oxides prepared in different ways gave definitely different effects.

In the case of aquo salt, the mechanism of the reaction and its relation to the medium is clear in at least one respect. The inhibition reactions which we have previously described are color tests with which we can test the intensities and selectivities of the force fields of the sixth "free" coördination valence. However, they are not as sensitive as the biological reaction just described. Groups with strong affinity to divalent iron in a complex (NH, SH, CN, NO, CO, etc.) block the free valence on entry into the complex and inactivate it, either wholly or partly. We have already mentioned that potassium ferrocyanide is entirely inactive and becomes activated on irradiation. Activation by means of light can also be biocatalytically demonstrated, for it is only irradiated potassium ferrocyanide that acts as an X mechanism.

These considerations give us a simple explanation of the conflicting results in the Webster-Baudisch experiments. If the bouillon contains substances which "block" the coördination valences there is no X action because the catalytic (autoxidative, peroxidative) action of the complex is either greatly diminished or entirely destroyed. When aquo salt is to be used as an X factor in bouillon cultures one must, at the very least, test its catalase action in the medium, as can be done by the Dubos-Baudisch method.

In applying this new information to the iron oxides, the methods were developed so that we could use synthetic media as well as bouillon. With these, also, we could demonstrate the selective action of certain oxides as X mechanisms. Previously it had been impossible to predict definitely biological activity by any physical or chemical tests.

These biological experiments have considerably clarified the complex X mechanism. The importance of catalase activity in iron oxides has come definitely to the fore and we must strive to prepare pure oxides that will be highly catalatic but not peroxidative. The previous emphasis on crystal structure retreats to the background. We now emphasize surface topography as it is determined by the pores of the clusters and by the degree of dispersion of the submicroscopic particles. It has become evident that

we are really only at the beginning of a knowledge of biologically active iron oxides. Our selection of  $\gamma$ - and  $\alpha$ -oxides as the "active" and "inactive" ones, respectively, was, in fact, only the key with which we opened up a very wide field of investigation.

Many investigators are engaged on the problem of structure in oxides prepared in various ways, and new facts and relations are constantly being published. For example, the structures of simple oxides have been considered in terms analogous to those employed in the classical valence theory of organic chemistry (Krause) (54). Along with efforts such as this we have the catalytic tests, the x-ray analyses, and the magnetic measurements.

It would be highly advisable and worth while to select the most characteristic types of iron oxides and use them in such biological experiments as have been described in this paper. The iron cyanide complexes are also susceptible to further development, as far as biological activity is concerned, as has recently been shown by Margarethe Lwoff (58) at the Pasteur Institute.

#### REFERENCES FOR TABLE 1

- (1) ANGELI-ARNDT: Sauerstoffhaltige Verbindungen des Stickstoffs. Samml. chem. chem.-tech. Vortr. 1908, XIII.
- (2) Baly, Edward Charles Cyril, Heilbron, Isidor Morris, and Stern, Harold Jacob: Trans. Chem. Soc. **123**, 185 (1923).
- (3) BAUDISCH, O.: Ber. **44**, 1009 (1911). Über Nitrat- und Nitrit-Assimilation. I.
- (4) BAUDISCH, O.: Schweiz. Naturforsch. Ges., 1911. Über Nitrat- und Nitrit-Assimilation. II.
- (5) BAUDISCH, O.: Zentr. Bakt. Parasitenk. **32**, No. 20/25 (1912). Über Nitrat- und Nitrit-Assimilation und über eine neue Hypothese der Bildung von Vorstufen der Eiweisskörper in den Pflanzen.
- (6) BAUDISCH, O., AND MAYER, ERWIN: Ber. **45**, 1771 (1912). Lichtchemische Vorlesungsversuche von pflanzenphysiologischen Interesse (III. lichtchemische Mitteilung).
- (7) BAUDISCH, O.: Ber. **45**, 2879 (1912). Über Nitrat- und Nitrit-Assimilation. IV. Eine Erwiderung an Herrn Oskar Loew.
- (8) BAUDISCH, O., AND MAYER, ERWIN: Ber. **46**, 115 (1913). Über Nitrat- und Nitrit-Assimilation.
- (9) BAUDISCH, O.: Vierteljahrsschr. Naturforsch. Ges. Zürich **59**, 1 (1913). Über das chemische Verhalten anorganischer stickstoffhaltiger Pflanzennahrungsstoffe gegenüber dem Sonnenlicht.
- (10) BAUDISCH, O.: Z. angew. Chem. **26**, 612 (1913). Über Nitrat- und Nitrit-Assimilation.
- (11) BAUDISCH, O.: Jahrb. Photogr. Reprod. Technik (1913). Über das chemische Verhalten anorganischer stickstoffhaltiger Pflanzennahrungsstoffe gegenüber dem Sonnenlicht.
- (12) BAUDISCH, O.: Naturwissenschaften **2**, 199, 229 (1914). Zur Frage der Assimilation anorganischer stickstoffhaltiger Verbindungen in dem Pflanzen.
- (13) BAUDISCH, O., AND MAYER, E.: Z. physiol. Chem. **89**, 176 (1914). Photochemische Studien zur Nitrat- und Nitrit-Assimilation.

- (14) BAUDISCH, O.: Jahrb. Photogr. Reprod. Technik (1914). Über chemische Lichtwirkung auf grossen Bergeshöhen.
- (15) BAUDISCH, O.: Ber. **49**, 1148 (1916). Über Nitrat- und Nitrit-Assimilation. VIII. Über Cholera. I.
- (16) BAUDISCH, O.: Ber. **49**, 1176 (1916). Über Nitrat- und Nitrit-Assimilation. IX.
- (17) BAUDISCH, O., AND KLINGER, G.: Ber. **49**, 1167 (1916). Über Nitrat- und Nitrit-Assimilation. X.
- (18) BAUDISCH, O.: Ber. **49**, 1176 (1916). Über Nitrat- und Nitrit-Assimilation. XI.
- (19) BAUDISCH, O.: Ber. **50**, 652 (1917). Über Nitrat- und Nitrit-Assimilation. XII. O. Loew nochmal zur Erwiderung.
- (20) BAUDISCH, O.: Ber. **51**, 793 (1918). Über Nitrat- und Nitrit-Assimilation. XIII. Eisen und Sauerstoff als notwendige Agenzien für die Reduktion von Alkalinitriten mit autoxydablen Verbindungen.
- (21) BAUDISCH, O.: Ber. **52**, 40 (1919). Über Nitrat- und Nitrit-Assimilation. XIV.
- (22) BAUDISCH, O.: Ber. **52**, 40 (1919). Über Nitrat- und Nitrit-Assimilation. XV.
- (23) BAUDISCH, O., AND MAYER, P.: Biochem. Z. **107**, 1 (1920). Studien über die Reduktion der Nitrite und Nitrats. Eine neue Methode zur quantitativen Bestimmung dieser Salze.
- (24) BAUDISCH, O.: J. Biol. Chem. **48**, 489 (1921). The Mechanism of Reduction of Nitrates and Nitrites in Processes of Assimilation.
- (25) BHARGANO, L. N., AND DHAR, N. R.: J. Indian Chem. Soc. **10**, 453 (1933).
- (26) HANTZSCH, A.: Ber. **66**, 1566 (1930).
- (27) HARTEK, PAUL: Ber. **66**, 423 (1933).
- (28) VON NAGEL, A.: Z. Elektrochem. **36**, 754 (1930).
- (29) VORLÄNDER, D., AND GORDES, WERNER: Ber. **64**, 1776 (1931). Die Allgegenwart von Stickoxyden.
- (30) ZINTEL, E., AND HARDER, A.: Ber. **66**, 760 (1933).

## REFERENCES

- (1) ALBRECHT, W. H.: Ber. **62**, 1475 (1929).
- (2) ANGELI-ARNDT: Samml. chem. chem.-tech. Vortr. 1908, XIII.
- (3) AVERY, O. T.: J. Exptl. Med. **34**, 38-40 (1921).
- (4) BAMBERGER, E.: Ber. **33**, 1781 (1900).
- (5) BAMBERGER, E.: Ber. **33**, 1939 (1900).
- (6) BAMBERGER, E., AND BAUDISCH, O.: Ber. **42**, 3568 (1909).
- (7) BARKER: Z. Krist. **45**, 1 (1908).
- (8) BAUDISCH, O.: Chem.-Ztg. **33**, 1298 (1909).
- (9) BAUDISCH, O.: Chem.-Ztg. **35**, 913 (1911).
- (10) BAUDISCH, O.: Ber. **49**, 172 (1916); **50**, 334 (1917).
- (11) BAUDISCH, O.: Naturwissenschaften **5**, 289 (1917). Innere Komplexsalz in ihrer Beziehung zur physiologischen Chemie.
- (12) BAUDISCH, O.: Ber. **53**, 413 (1920).
- (13) BAUDISCH, O.: Ber. **54**, 413 (1921).
- (14) BAUDISCH, O.: Ber. **62**, 2699 (1929). Die Bedeutung von Austausch oder Verdrängungsreaktionen bei katalytischen Vorgängen.
- (15) BAUDISCH, O.: Biochem. Z. **232**, 35 (1931). Über den Einfluss koordinative gebundener Gruppen auf die Eigenschaften des zentralen Eisenatoms in Eisencyanverbindungen.
- (16) BAUDISCH, O.: Biochem. Z. **245**, 265 (1932).

- (17) BAUDISCH, O.: *Biochem. Z.* **258**, 69 (1933).
- (18) BAUDISCH, O.: *Aktives Eisen*. Carl Oppenheimer Handbuch der Biochemie des Menschen und der Tiere, 2te Auflage, Ergänzung zu Bd. I-III des Hauptwerks, p. 749. Gustav Fischer, Jena (1933).
- (19) BAUDISCH, O., AND ALBRECHT, W. H.: *Naturwissenschaften* **20**, 639 (1932).
- (20) BAUDISCH, O., AND ALBRECHT, W. H.: *J. Am. Chem. Soc.* **54**, 943 (1932).
- (21) BAUDISCH, O., AND DAVIDSON, DAVID: *J. Biol. Chem.* **71**, 501 (1927).
- (22) BAUDISCH, O., AND DUBOS, R.: *Biochem. Z.* **245**, 278 (1932).
- (23) BAUDISCH, O., AND FÜRST, ROSE: *Ber.* **48**, 1665 (1915).
- (24) BAUDISCH, O., GUREWITSCH, H., AND ROTHSCHILD, S.: *Ber.* **49**, 180 (1916).
- (25) BAUDISCH, O., AND KARZEFF, N.: *Ber.* **45**, 1164 (1912).
- (26) BAUDISCH, O., AND KING, VICTOR L.: *Ind. Eng. Chem.* **3**, 629 (1911).
- (27) BAUDISCH, O., PISTOR, A. E., AND SILBERBLATT, B.: *Ber.* **49**, 191 (1916).
- (28) BAUDISCH, O., AND ROM, HERMANN: *Ber.* **49**, 203 (1916).
- (29) BAUDISCH, O., AND ROTHSCHILD, S.: *Ber.* **48**, 1660 (1915).
- (30) BAUDISCH, O., AND WELO, L. A.: *J. Biol. Chem.* **61**, 261 (1924).
- (31) BAUDISCH, O., AND WELO, L. A.: *Naturwissenschaften* **21**, 659 (1933).
- (32) BILTZ, W.: *Z. anorg. allgem. Chem.* **145**, 63 (1925). Résumé of references.
- (33) BILTZ, W.: *Z. anorg. allgem. Chem.* **164**, 245 (1927).
- (34) BILTZ, W.: *Z. anorg. allgem. chem.* **164**, 245 (1927); **166**, 351 (1927).
- (35) BILTZ, W., AND HÜTTIG, G.: *Z. anorg. allgem. Chem.* **114**, 241 (1920).
- (36) BÖHM, I.: *Z. Krist.* **68**, 567 (1928).
- (37) CAMBI, L.: *Chem. Zentr.* **1914**, II, 1100.
- (38) CAMBI, L., AND SZEGÖ, L.: *Ber.* **66**, 656 (1933).
- (39) COSTACHESCU, N., AND SPACU, G.: *Chem. Zentr.* **1912**, II, 1972.
- (40) DAVIDSON, D.: *J. Am. Chem. Soc.* **50**, 2622 (1928).
- (41) EICHHOLTZ, F., AND BERG: *Biochem. J.* **225**, 353 (1930).
- (42) EPHRAIM, F.: *Ber.* **50**, 1069 (1917); **51**, 644 (1918).
- (43) FEIGL, F., AND BONDI, A.: *Ber.* **64**, 2819 (1931).
- (44) GRIMM, H.: *Naturwissenschaften* **17**, 536 (1929). Isostere Molekeln.
- (45) GROSSMANN AND HÜNSELER: *Z. anorg. allgem. Chem.* **46**, 370 (1905).
- (46) HAUROWITZ, F., AND KITTEL, H.: *Ber.* **66**, 1046 (1933).
- (47) HENSCHEL, H., AND RINNE, FR.: *Ber. Sachs. Ges. Wiss. math.-phys. Klasse* **79**, 17, 1 (1927).
- (48) HIEBER, S.: *Z. Elektrochem.* **39**, 27 (1933); *Ber.* **60**, 2300 (1927); *Z. anorg. allgem. Chem.* **186**, 102 (1930); **190**, 197 (1930).
- (49) HOFMANN, K. A.: *Ann.* **312**, 1 (1900).
- (50) JOHNSON, T. B., AND BAUDISCH, O.: *J. Am. Chem. Soc.* **43**, 2670 (1921).
- (51) KLAUS, F., AND BAUDISCH, O.: *Ber.* **51**, 1036 (1918).
- (52) KLEMM, W., JACOBI, H., AND TILK, W.: *Z. anorg. allgem. Chem.* **201**, 1 (1931).
- (53) KLEMM, W.: *Z. anorg. allgem. Chem.* **201**, 1 (1931); *Z. angew. Chem.* **190**, 193 (1930).
- (54) KRAUSE, A.: *Z. anorg. allgem. Chem.* **208**, 282 (1932).
- (55) LANGMUIR, I.: *J. Am. Chem. Soc.* **41**, 1573 (1919).
- (56) LANGMUIR, I., AND HÜCKEL, W.: *Z. Elektrochem.* **27**, 305 (1921).
- (57) LEY, H.: *Die Beziehungen zwischen Farbe und Konstitution*. S. Hirzel, Leipzig (1911).
- (58) LWOFF, M.: *Compt. rend. soc. biol.* **107**, 1438 (1931); *Ann. inst. Pasteur* **50**, 1, 55, 707 (1933).
- (59) MAGNUS, A.: *Z. anorg. allgem. Chem.* **124**, 293 (1922).

- (60) MEISEL, K., AND TIEDJE, W.: *Z. anorg. allgem. Chem.* **164**, 223 (1927).
- (61) MICHAELIS, L., AND SMYTHE, C. V.: *J. Biol. Chem.* **94**, 329 (1931).
- (62) MÜLLER, E.: *Z. Elektrochem.* **31**, 43 (1925).
- (63) PASCAL: *Rev. gén. sci.*, July, 1923. *Diamagnetism et Constitution Chimique.*
- (64) PFALTZ, M. H.: *J. Am. Chem. Soc.* **45**, 2980 (1923).
- (65) PFALTZ, M. H., AND BAUDISCH, O.: *J. Am. Chem. Soc.* **45**, 2972 (1923).
- (66) PFEIFFER, P.: *Ber.* **63**, 1811 (1930).
- (67) RABINOWITSCH, E.: *Z. Elektrochem.* **39**, 702 (1933).
- (68) RAIKOWA, THEODORA P.: *Ber.* **62**, 1626, 2142 (1929).
- (69) RAY, R.: *Z. anorg. allgem. Chem.* **174**, 189 (1928).
- (70) ROSENHEIM, A.: *Z. anorg. allgem. Chem.* **216**, 173 (1934).
- (71) ROSENHEIM, A., ROEHRICH, E., AND TREWENDT, L.: *Z. anorg. allgem. Chem.* **207**, 107 (1932).
- (72) SCHWARZ, ROBERT, AND BASS, LAWRENCE W.: *The Chemistry of the Inorganic Complex Compounds.* John Wiley and Sons, Inc., New York (1923).
- (73) SPACU, G.: *Z. anorg. allgem. Chem.* **216**, 165 (1934).
- (74) STELLING, O.: *Z. physik. Chem.* **16B**, 317 (1932).
- (75) THILO, ERICH: *Die Valenz der Metalle Fe, Co, Ni, Cu und ihre Verbindungen mit Dioximen.* Samml. chem. chem.-tech. Vortr., neue Folge, Heft 13. Stuttgart (1932).
- (76) TRAUBE, W., AND WOLFF, W.: *Ber.* **60**, 43 (1927).
- (77) VAN ARKEL AND DE BOER: *Chemische Bindung als elektrostatische Erscheinung*, p. 161. Translated by Li Klemm and Wilhelm Klemm. S. Hirzel, Leipzig (1931).
- (78) WEBSTER, L. T.: *Proc. Soc. Exptl. Biol.* **22**, 139 (1924).
- (79) WEBSTER, L. T., AND BAUDISCH, O.: *J. Exptl. Med.* **42**, 473 (1925).
- (80) WEINLAND, R.: *Einführung in die Chemie der Komplexverbindungen*, 2te Auflage. Ferdinand Enke, Stuttgart (1924).
- (81) WELO, LARS A.: *Phil. Mag.* **6**, 481 (1928).
- (82) WELO, L. A., AND BAUDISCH, O.: *Nature* **116**, 606 (1925).
- (83) WELO, L. A., AND BAUDISCH, O.: *Phil. Mag.* **6**, 50, 399 (1925).
- (84) WERNER, A.: *Neuere Anschauungen auf dem Gebiete der anorganischen Chemie*, 5te Auflage, neu bearbeitet von P. Pfeiffer. Braunschweig (1923).
- (85) WYCKOFF, R. W. G., AND McCUTCHEON, TH. P.: *Am. J. Sci.* **13**, 223 (1927).





## ACTIVE IRON. II

### RELATIONSHIPS AMONG THE OXIDE HYDRATES AND OXIDES OF IRON AND SOME OF THEIR PROPERTIES

LARS A. WELO AND OSKAR BAUDISCH

*Sloane Physics and Sterling Chemistry Laboratories, Yale University, New  
Haven, Connecticut*

*Received April 3, 1934*

#### I. INTRODUCTION

A study of the extensive literature (2, 32) on synthetic oxide hydrates and oxides of iron, both ferrous and ferric, suggests a simple way of classifying them and of exhibiting the relationships among the various members. This is to express their generic relations to the iron compounds from which they were formed and to specify the chemical operations that were used.

The experimenters in the field of our survey started with either a ferrous salt, a mixture of ferrous and ferric salts, or a ferric salt, and carried out at least one of the following operations: (1) Precipitation with an alkali. In the case of the ferrous hydrates and oxides, this precipitation was carried out under oxygen-free conditions. (2) Oxidation with either molecular oxygen or with some oxidizing agent containing oxygen. (3) Dissociation. (4) Hydrolysis. (5) Heating. (6) Aging. By aging is meant the changes that take place spontaneously. Unless otherwise specified the temperatures at which aging occurs are those of the room. The distinction between heating and aging is generally arbitrary, since they differ only as to rate. The justification for considering aging as a chemical operation is the fact that the paths and the rates can, in general, be modified by changing the chemical environment.

The general scheme of classification that has been indicated has led us to construct the diagram of figure 1. We shall let this diagram serve as an outline throughout our discussion of the synthetic hydrates and oxides. Necessarily, the physical and chemical properties must be considered member by member. It is hoped that frequent reference to the diagram will maintain a certain coherence among the various parts of this review.

It was not deemed advisable to complicate the diagram by indicating special methods of preparation which do not fit readily into the general scheme. None of these appears to affect the indicated relationships. To

indicate, for example, that the ferric oxides may be reduced to  $\text{Fe}_3\text{O}_4$  or  $\text{FeO}$  would not add to the value of the diagram as a generic scheme. If pertinent to the main discussion, special methods will be described during the discussion of the member that is concerned.

It must not be assumed that we regard this diagram as a permanent representation of the relationships between the hydrates and oxides of iron, and that the only future changes will be to add new members or new series. In our view, the existence of some members and the paths of some changes are doubtful. The diagram is intended only to summarize prevailing opinions among the more active workers in this field. We believe that a

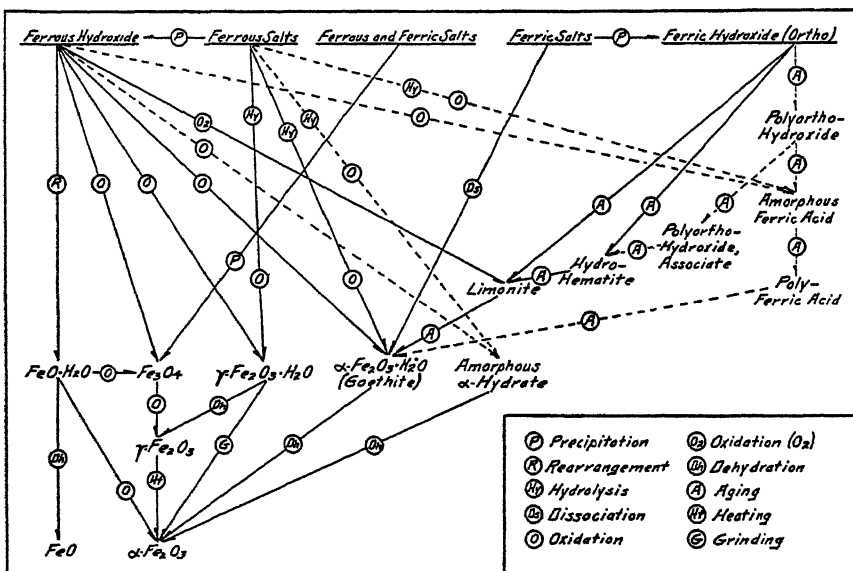


FIG. 1. RELATIONSHIPS AMONG THE OXIDE HYDRATES AND OXIDES OF IRON

review does not serve its full purpose if it records only positive and well-established knowledge. A review should, rather, emphasize the existence of uncertainties. It should present a field as a growing and developing one and should allow free presentation of the newer results obtained by newer methods. New results must be tested and weighed by numerous workers before final acceptance. Time and further studies, alone, will bring truth.

Finally, it will be apparent that this review will reflect, to a considerable extent, our own particular interests. For this we offer no apologies. Lack of space compels limitation. We shall not describe all properties nor list all physical and chemical constants. We shall limit ourselves to significant

phenomena that occur in the paths of change as they may be followed by the dehydration isobars, the x-ray spectra, the magnetic properties, the specific heats, the silver analysis, and the catalytic and chemical activities. We consider that these are, precisely, the experimental methods and data that have led to the progress of recent years. And it is with the more recent results that this review will be mainly concerned.

## II. FERROUS HYDROXIDE, OR FERROUS OXIDE MONOHYDRATE

The snow-white flocculent substance which is precipitated when an alkali is added to a solution of a ferrous salt is perhaps the least known and understood of the hydrates of iron. Indeed, few chemists have ever

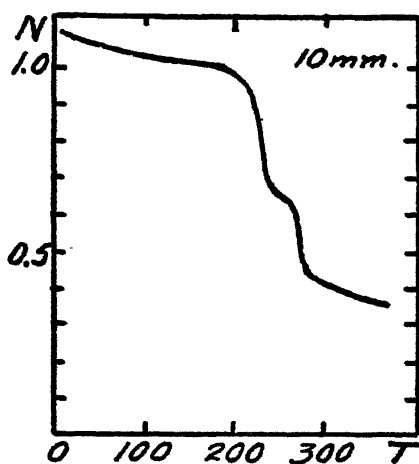


FIG. 2. DEHYDRATION ISOBAR FOR  $\text{FeO} \cdot \text{H}_2\text{O}$   
(Hüttig and Möldner)

seen it in its snow-white state because of its sensitiveness to minute traces of oxygen. It is necessary to boil both the salt and alkali solutions previous to mixing in order to eliminate dissolved oxygen, and the apparatus (15) should be arranged in such a way that all operations and subsequent observations may be made in the presence of some oxygen-free neutral gas such as nitrogen or hydrogen. Another effective way of obtaining oxygen-free conditions is to boil the solutions separately while they are covered with a thick layer of vaseline or castor oil and then to pour the contents of one vessel, both solution and covering material, into the other. Ferrous hydroxide is usually obtained with a green color which turns black or red according to the amount of oxygen that is admitted. The difficulty of insuring strictly oxygen-free conditions has discouraged chemists from a

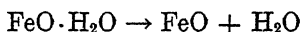
study of its properties. They have been interested, rather, in its strong reducing properties, and in the higher oxides that are so easily formed.

The first definite information as to chemical structure was obtained by Hüttig and Möldner (49), who proved the existence of the monohydrate,  $\text{FeO} \cdot \text{H}_2\text{O}$ , by means of the dehydration isobar shown in figure 2. They recognized and discussed the irregularity in the range 0.65 to 0.40  $\text{H}_2\text{O}$ , but its presence does not vitiate the main conclusion as to the existence of the monohydrate.

Hüttig and Möldner did not analyze the x-ray spectra of the monohydrate in detail, but were satisfied with the conclusion that they differed from the spectra of  $\text{FeO}$ ,  $\text{Fe}_3\text{O}_4$ ,  $\alpha$ - or  $\gamma$ - $\text{Fe}_2\text{O}_3 \cdot \text{H}_2\text{O}$ , and  $\alpha$ - or  $\gamma$ - $\text{Fe}_2\text{O}_3$ . They were probably identical with the spectra of Natta and Cassaza (71), from which it was determined that ferrous oxide hydrate is rhombohedral with  $a = 3.24$ ,  $c = 4.47$ , and  $c/a = 1.38$ .

Ferrous oxide hydrate is known to be paramagnetic. Quartaroli's (76) experimental methods hardly excluded oxygen and the formation of higher oxides, especially magnetite, so that one must favor his smallest susceptibility value,  $\chi = +20 \times 10^{-6}$ .

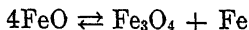
Hüttig and Möldner observed that small amounts of hydrogen were formed when they dehydrated ferrous oxide hydrate. They assumed that in addition to the main reaction



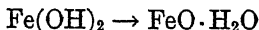
there was another



The black dehydration product that they describe was, therefore, ferrous oxide with an admixture of magnetite,  $\text{Fe}_3\text{O}_4$ . This is in accord with the result of Chaudron and Forestier (22) that ferrous oxide is not stable below  $570^\circ\text{C}$ . There should also be some free iron because of the reaction



Using a thermodynamic argument, Hüttig and Möldner make it appear probable that ferrous oxide hydrate is not stable at ordinary temperatures and that some slow change does occur. Whether it is this change beyond ferrous oxide hydrate or whether it is a previous change implied in a rearrangement like



is not known, but Baudisch and Welo (17) have shown that a spontaneous change does take place. They found that a hydrate that had been aged

for four days in a neutral solution, in the absence of air, yielded  $\alpha$ -ferric oxide when oxygen was admitted, together with a small amount of  $\gamma$ -ferric oxide, as indicated by a weak ferromagnetism. This is entirely opposite to what is observed when oxidation is carried out shortly after precipitation. On oxidation freshly precipitated ferrous oxide hydrate yields  $\text{Fe}_3\text{O}_4$  in alkaline solution,  $\alpha\text{-Fe}_2\text{O}_3 \cdot \text{H}_2\text{O}$  in acid or neutral solution and, in certain circumstances,  $\gamma\text{-Fe}_2\text{O}_3 \cdot \text{H}_2\text{O}$ . This experiment has significance in connection with the difference between fresh and aged ferrous oxide hydrate as a reducing agent in the reduction of alkali nitrates (7, 13, 14, 15).

It is only in the presence of molecular oxygen (air) that freshly precipitated ferrous hydroxide can reduce alkali nitrates to nitrites or oxidize oxidizable substances that may be present, such as lactic acid to pyruvic acid or uracil to isobarbituric acid. In the absence of oxygen, either with or without nitrate, the hydroxide remains snow-white even when left to stand for years, showing that it does not dissociate water into hydrogen and oxygen. However, after long standing or aging under anaerobic conditions, ferrous hydroxide is unable either to reduce nitrates or to oxidize lactic acid and uracil. But it remains autoxidizable and may be completely oxidized to ferric hydroxide.

This peculiar behavior of ferrous hydroxide raises three questions: (1) What happens to ferrous hydroxide when it ages in the absence of air? (2) Why is oxygen necessary in order that ferrous hydroxide may reduce nitrates? More specifically, why is oxygen necessary for the production of activated hydrogen? (3) Whence comes the activated oxygen which is able to oxidize lactic acid to pyruvic acid or uracil to isobarbituric acid?

The fact that oxidation of aged ferrous hydroxide yields an oxide and neither of the hydrates points to the conclusion that the aging under anaerobic conditions involves the formation of a compound with water bound in another way.

Answers to the second and third questions are suggested by the work of Goldberg (33), who found that ferrous hydroxide dissociates water to liberate hydrogen and oxygen when the temperature is  $200^\circ\text{C}$ . or more. It is known that even at room temperature, in the presence of oxygen, ferrous hydroxide can reduce nitrates or oxidize certain organic compounds. Hence it appears that if oxygen is present, water is dissociated by ferrous hydroxide even at room temperature, according to the scheme:  $\text{H}_2\text{O} \rightarrow \text{H} + \text{OH}$  ( $\text{H}_2\text{O}_2$ ).

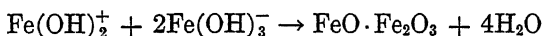
Since the aged ferrous hydroxide can bring about neither oxidation nor reduction, and since the outstanding difference between the aged and the fresh ferrous hydroxide is the absence or presence of this water of hydration, it follows that the water which is dissociated into activated hydrogen and oxygen is this labile water. It is neither the adsorbed nor the solvent

water. Ferrous hydroxide which has been aged for a sufficient length of time and has no labile water left has, therefore, lost the power to reduce nitrates or to oxidize uracil and lactic acid.

### III. THE FORMATION OF MAGNETITE, $\text{Fe}_3\text{O}_4$

In order to obtain the partly oxidized magnetite and not the fully oxidized  $\text{Fe}(\text{OH})_3$  on oxidation of  $\text{Fe}(\text{OH})_2$  (we retain the conventional notation), the rate of oxygen supply must be low and the liquid medium must be alkaline. The latter condition was emphasized by Miyamoto (69) when he found that high hydroxyl-ion concentration retarded oxidation, which is apparently in contradiction with the well-known fact that ferrous hydroxide is a stronger reducing agent in an alkaline medium. Krause (53) and Deiss and Schickorr (26) have given the explanation and resolved the contradiction.

They regard the full oxidation of ferrous hydroxide to ferric hydroxide as the primary process. According to Krause the isoelectric point of ferric hydroxide, when formed by oxidation of ferrous hydroxide, is at  $\text{pH} = 5.2$ . It finds itself in an alkaline medium and immediately assumes an acid character. It then combines with the still unoxidized, strongly basic ferrous hydroxide, of isoelectric point  $\text{pH} > 11.5$ , to form ferrous ferrite according to the reaction:

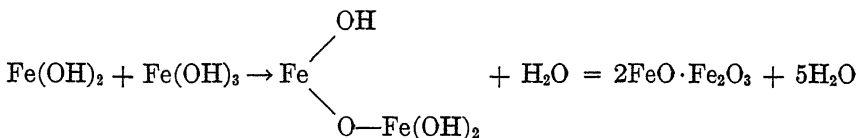


Ferrous ferrite does not oxidize further under these conditions.

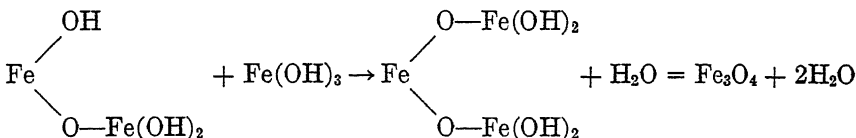
When the rate of oxygen supply is low enough there is sufficient time for reaction between ferrous and ferric hydroxides to form ferrous ferrite. At boiling temperature, particularly, the formation of ferrous ferrite is favored because the rate of combination is increased and at the same time the supply of oxygen is diminished. On the other hand, when the supply of oxygen is plentiful, and especially at low temperatures when the rate of oxidation is high as compared with the rate of combination of ferrous hydroxide and ferric hydroxide, we will have nearly complete oxidation to ferric hydroxide. This view of the matter explains very well the fact that a favored way of preparing magnetite from ferrous hydroxide is to work at boiling temperatures. And it explains the interesting fact that a piece of iron kept at the bottom of a vessel of water becomes covered with a black rust, while one suspended just beneath the surface acquires a reddish brown coating.

While the explanation that has been given suffices to account for the formation of magnetite instead of ferric hydroxide, it does not account for the green intermediate product that is always observed. Deiss and Schickorr noticed that the green substance is formed after a momentary appearance

of the red ferric hydroxide and assumed that the green compound is also a combination of ferric hydroxide with unoxidized ferrous hydroxide. A possible combination is:



The formation of magnetite would then involve the combination of this green substance with newly formed ferric hydroxide, according to the scheme



Whether this path or the more direct one is followed, the primary process is the oxidation of ferrous hydroxide to ferric hydroxide, followed by reaction between these two kinds of hydroxides. The effect of varying the rate of oxidation would be the same. Deiss and Schickorr were unable, because of experimental difficulties, to secure analytical proof that the green compound had the composition  $2\text{FeO} \cdot \text{Fe}_2\text{O}_3$ .

Essentially the same principle of the reaction between ferrous hydroxide and ferric hydroxide is involved in the preparation of magnetite by precipitation of an equimolar solution of ferrous and ferric salts with an alkali (61, 104). Ferric hydroxide is already present so that no oxidation is necessary. While the crystal structures and the magnetic properties of magnetite obtained in this way correspond to those of magnetite obtained by slow oxidation of ferrous hydroxide, there are marked differences in detail which follow from the altered conditions of precipitation.

These differences arise from the different solubilities of ferrous hydroxide and ferric hydroxide and from the fact that when ferrous hydroxide is oxidized there are two opportunities for the crystallization process to operate, whereas in the simultaneous precipitation of ferrous hydroxide and ferric hydroxide the crystallization process operates only once and then under very unfavorable conditions.

Ferrous hydroxide is not extremely insoluble. The accumulation velocity (the *Häufungsgeschwindigkeit* of Haber (36)) is also not extremely high so that there is time for the formation of good crystals of considerable size; that is, the rate of crystallization (*Ordnungsgeschwindigkeit*) is not exceeded by the accumulation rate. The oxidation which follows involves



recrystallization and is slow, so that the crystals of this magnetite are large and well developed, as shown by the sharpness of the x-ray lines, and are free from strain, since the magnetic properties do not alter greatly on annealing.

Ferric hydroxide, on the contrary, is extremely insoluble in water, and the rate of accumulation is high, since it is proportional to the density excess over the solubility. It exceeds the rate of crystallization and the crystals, if any, are imperfect. Also, according to the precipitation laws they should be small. The x-ray lines of this magnetite, formed by simultaneous precipitation of ferrous hydroxide and ferric hydroxide, are diffuse, showing that the crystals are either of small size, or imperfect, or both. The magnetic properties alter greatly on annealing.

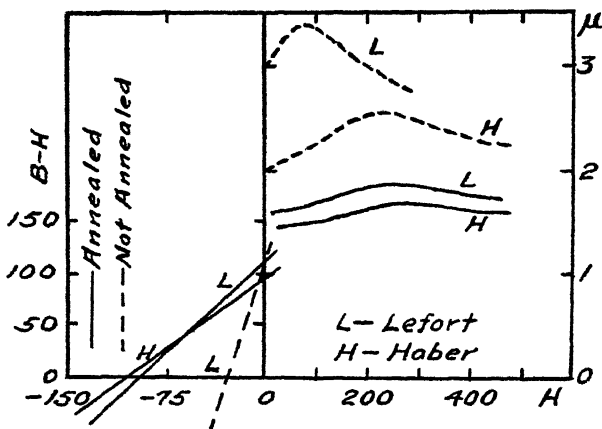


FIG. 3. EFFECT OF ANNEALING AT 1000°C. ON PERMEABILITIES AND COERCIVITIES OF MAGNETITES

Figure 3 shows the permeability curves and portions of the hysteresis loops for two magnetites studied by Welo and Baudisch (104). The intersections of the hysteresis loops with the axes give the remanences and the coercive forces. "Lefort's" refers to the magnetite obtained by simultaneous precipitation of ferrous hydroxide and ferric hydroxide. "Haber's" refers to the magnetite formed by oxidation of ferrous hydroxide precipitated with ammonia in the presence of potassium nitrate at boiling temperatures. Haber's undergoes very little change on annealing at 1000°C., from which it is inferred that the crystals were initially well formed. Lefort's is markedly changed on annealing, and its magnetic properties approach those of Haber's. Lefort's was certainly imperfectly crystallized. Results since obtained, which will be considered later, make it probable that smallness of crystal size also plays a part.

The special place which Lefort's magnetite occupies as the most suitable material for the preparation of "active iron oxide" is due to its small crystals and to their poor quality as brought about by the precipitation conditions. Also, the unlike chemical characteristics of ferric hydroxide obtained by oxidizing ferrous hydroxide, and of ferric hydroxide obtained directly by precipitation from a ferric salt solution may persist in the magnetites formed from them. To this possibility, serious consideration has not yet been given. Krause and Pilawski (58) assign different structural formulas for these hydrates (the meta- and ortho-hydroxides considered in the last section of this review) on the basis of their silver analysis. It is known that the isoelectric point of the former is at  $\text{pH} = 5.2$  and of the latter at  $\text{pH} = 7.7$  (53).

Magnetite has been prepared by oxidizing ferrous hydroxide with nitric oxide (104). A magnetite that is free of impurities other than 0.001 per cent of carbon, and which has the very rare property of having exactly the theoretical percentage of ferrous oxide, 31 per cent, has been obtained by burning iron carbonyl in a controlled air supply (68). The group of French workers (3, 22, 30, 43) who have devoted themselves to thermomagnetic analyses have made their magnetites by reducing  $\alpha$ -ferric oxide with hydrogen in the presence of water vapor at temperatures of 300 to 600°C. Another way to reduce ferric oxides to magnetite is to leave them for a short time in molten sodium acetate (105). To describe, or even mention, all methods of preparing magnetite would lead us too far afield.

The various formulations of precipitated magnetite found in the literature, such as  $2\text{FeO} \cdot 3\text{Fe}_2\text{O}_3$ ,  $4\text{FeO} \cdot \text{Fe}_2\text{O}_3$ , etc., are improbable. The possibility of forming mixtures of  $\text{Fe}_3\text{O}_4 = \text{FeO} \cdot \text{Fe}_2\text{O}_3$  with  $\text{FeO} \cdot \text{H}_2\text{O}$  and ferromagnetic ferric oxide is inherent in the precipitation and oxidation method. X-ray analyses have always given the same set of lines regardless of the procedure in preparation (104). The x-ray lines yield the same spacings of the lattice planes as are found in natural magnetite and in magnetite derived by dry reduction of ferric oxide. An underoxidized magnetite studied by Hüttig and Möldner (49) showed lines of ferrous oxide hydrate. Lefort's magnetite is generally overoxidized, but the excess ferric oxide is ferromagnetic ferric oxide which, as will be seen, has the same x-ray spectrum as magnetite except for slight differences in relative intensities.

The x-ray analyses indicate, in addition, that there are no definite hydrates of magnetite. This conclusion is supported by the dehydration isobars obtained by Hüttig and Möldner (49), whose curves are shown in figure 4. Curve A is for a sample prepared by oxidizing ferrous hydroxide with oxygen ( $\text{O}_2$ ). It was slightly underoxidized with  $\text{Fe}:\text{O} = 1:1.289$ , instead of  $\text{Fe}:\text{O} = 1:1.333$  as in the theoretical ratio for  $\text{Fe}_3\text{O}_4$ . B is for a part of sample A which had been aged for 70 days in a sealed container at

room temperature. C is for a magnetite made by simultaneous precipitation of a mixture of ferrous and ferric chlorides giving  $\text{Fe}:\text{O} = 1:1.335$ . Curve D was obtained with an underoxidized product,  $\text{Fe}:\text{O} = 1:1.196$ . In curves A, B, and C there are no discontinuities and hence no indications of the existence of definite hydrates. A comparison of A and B shows the characteristic feature of the aging process in oxide hydrates containing water other than water of constitution. It is its loosening and gradual elimination. The discontinuity in curve D is due to the ferrous monohydrate whose dehydration isobar is shown in figure 2.

The crystal structures and magnetic properties of magnetite are so nearly like those of the ferromagnetic ferric oxide derived by oxidation of magnetite, hereafter called  $\gamma$ -oxide, that they will be considered together.

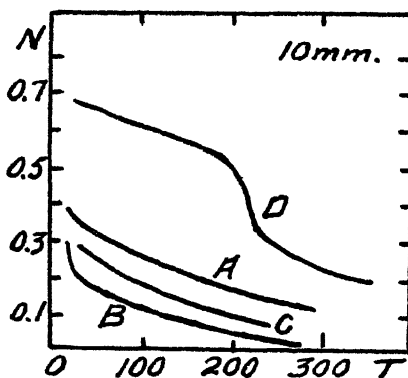


FIG. 4. DEHYDRATION ISOBARS FOR  $\text{Fe}_3\text{O}_4 \cdot x\text{H}_2\text{O}$  (Hüttig and Möldner)

#### IV. $\gamma$ -OXIDE

To describe the behavior of precipitated magnetite as it is oxidized at a low temperature, the magnetic and crystal structure transformations that occur at higher temperatures, and the crystal structure problem that arises, we shall quote in large part from our own paper (102).

"Generally speaking, the facts observed on heating the oxide in a stream of oxygen are as follows:—The color of the oxide changes suddenly from black to a dark red at about  $220^\circ\text{C}$ ., the oxide remains magnetic, although the iron has been completely converted into the ferric form so that the composition has become  $\text{Fe}_2\text{O}_3$ . The ferricyanide test showed not a trace of Prussian blue, and hence no ferrous iron after oxidation for a sufficient time at  $220^\circ\text{C}$ . A solution of the oxide in hydrochloric acid titrated with potassium permanganate likewise yielded no iron in the divalent form. All of this shows that  $\text{Fe}_3\text{O}_4$  may be completely oxidized to  $\text{Fe}_2\text{O}_3$  at a temperature of  $220^\circ\text{C}$ . without loss of the magnetic properties.

"On further heating, the oxide, now  $\text{Fe}_2\text{O}_3$ , remains magnetic until a temperature of  $550^\circ\text{C}$ . is reached. At this temperature a non-reversible transformation takes place, for the oxide becomes permanently non-magnetic. This transformation is obviously a pure temperature effect, since the oxidation was already complete at the lower temperature of  $220^\circ\text{C}$ . But to test the point a sample which had first been oxidized at  $220^\circ\text{C}$ . was heated in an atmosphere of nitrogen carefully freed from oxygen. It again became non-magnetic. The necessary condition, however, for the loss of magnetism at  $550^\circ\text{C}$ . is that the  $\text{Fe}_3\text{O}_4$  must first be oxidized to  $\text{Fe}_2\text{O}_3$  at some lower temperature. If  $\text{Fe}_3\text{O}_4$  is heated in nitrogen without a preliminary oxidation it is stable and retains its magnetic properties even if heated to  $800^\circ\text{C}$ . These

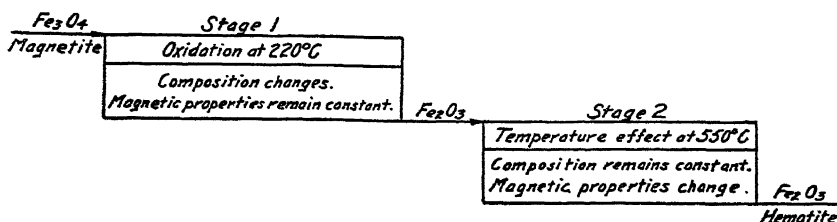


FIG. 5. TWO-STAGE TRANSFORMATION FROM MAGNETITE TO HEMATITE

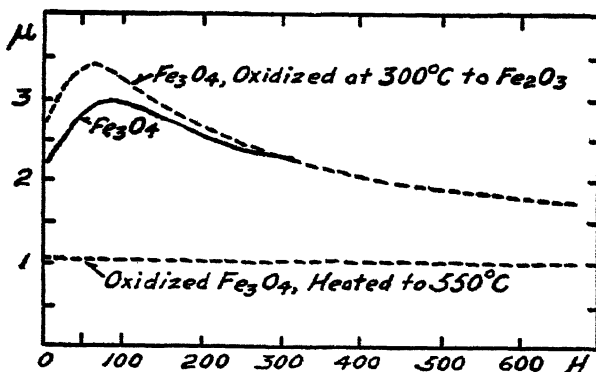


FIG. 6. PERMEABILITIES OF SYNTHETIC IRON OXIDES IN POWDERED FORMS

Densities of packing = 1.32 g. per cubic centimeter

observations are briefly summarized in the scheme shown in figure 1. (Figure 5 of this review.)

"It has been stated that the oxidation begins at  $220^\circ\text{C}$ . A more precise statement would be that the oxidation becomes sufficiently rapid at that temperature so that the change in color may be observed. It is recognized that oxidation may take place at a much lower temperature.

"The results of the magnetic measurements are shown in figure 2 (figure 6 of this review). Our qualitative observations are confirmed. The maximum permeability of the oxide after having passed the first stage of transformation is somewhat higher than that of the original  $\text{Fe}_3\text{O}_4$ , and there is shown a considerable shift of its position.

No reason is known for this. At higher fields the curves coincide, which accounts for our qualitative observation (made with high field intensities) that the magnetic properties remain unchanged on oxidation at a low temperature. After the second stage the permeability is seen to be nearly, but not quite, unity at all fields. The maximum residual permeability is 1.045, and is so small that the material would be considered non-magnetic towards rough qualitative tests.

"The crystal structures of magnetite and hematite are known to be different. Bragg (20) and Nishikawa (72) found for magnetite a diamond point structure and Davey (25) found that hematite is like corundum, which, in turn, resembles calcite. It was of interest to determine whether the change of one structure into the other takes place at stage I or at stage II. It seemed improbable that it changes at stage I on account of the constancy of the magnetic properties when the oxide passes through that stage. On the other hand, it was difficult to assume that it changes on stage II, for it would involve the assumption that  $\text{Fe}_2\text{O}_3$  has allotropic forms, and can assume either the form for magnetite or the calcite structure usually obtained for hematite.

"Diffraction patterns of the original  $\text{Fe}_3\text{O}_4$  and the resulting oxides after each stage of transformation were obtained. . . . The original  $\text{Fe}_3\text{O}_4$  and the  $\text{Fe}_2\text{O}_3$  which had passed through stage I gave identical patterns, and the spacings were the same to within one-tenth of one per cent. The  $\text{Fe}_2\text{O}_3$  which had been through stage II of the transformation, however, showed the pattern usually found for  $\text{Fe}_2\text{O}_3$ . The disappearance of the magnetic properties of the oxide is therefore associated with the change in structure taking place at stage II and is not connected with the change in composition. It is also evident that the atoms of iron and oxygen in  $\text{Fe}_2\text{O}_3$  have two alternative arrangements in the crystal. If  $\text{Fe}_2\text{O}_3$  is derived by oxidation of  $\text{Fe}_3\text{O}_4$  at a low temperature, the  $\text{Fe}_3\text{O}_4$  structure is retained, and only at a higher temperature of about  $550^\circ\text{C}$ . does this structure become unstable and change into the more stable one similar to calcite. The original  $\text{Fe}_3\text{O}_4$ , however, as has already been pointed out, remains stable up to temperatures of at least  $800^\circ\text{C}$ . if we prevent oxidation to  $\text{Fe}_2\text{O}_3$ .

"This result suggested the interesting problem of finding room for more oxygen in a crystal of  $\text{Fe}_3\text{O}_4$  without a change in the arrangement of the atoms already in place. . . .

"The unit cube contains eight molecules, so that  $\text{Fe}_3\text{O}_4$  may be written  $\text{Fe}_{24}\text{O}_{32}$ . After oxidation to  $\text{Fe}_2\text{O}_3$  the corresponding composition would be  $\text{Fe}_{24}\text{O}_{36}$ . We must find room for four additional oxygen atoms in the cube."

Miss Twenhofel (97) concluded that definite positions could not be assigned to the extra oxygen atoms, but Thewlis (93), after a thorough and critical study, has done this. He made photometric measurements on his x-ray films and found that the patterns of magnetite and of  $\gamma$ -ferric oxide, although very similar, are not identical. The reflections from the 400 and 440 planes are definitely more intense in  $\gamma$ -ferric oxide than in magnetite. His analysis involved the following special procedures in addition to those usually followed in crystal analysis.

(1) Determination of the space groups which can accommodate thirty-six oxygen atoms in the unit cube and at the same time allow twenty-four iron atoms to occupy the same positions as in magnetite. He found six such groups:  $O^6$ ,  $T_d^2$ ,  $T_d^1$ ,  $T^4$ ,  $T^2$ , and  $T^1$ .

(2) Decision as to the minimum permissible iron to oxygen and oxygen to oxygen distances. The smallest iron to oxygen distance in magnetite is the  $\text{Fe}^{++}$  to O distance of 1.88 A.U. (23) and the  $\text{Fe}^{++}$  and  $\text{Fe}^{+++}$  radii are, respectively, 0.78 A.U. and 0.63 A.U. (34). During oxidation  $\text{Fe}^{++}$  becomes  $\text{Fe}^{+++}$ , so that the least iron to oxygen distance in  $\gamma$ -ferric oxide can be shorter than in magnetite by 0.15 A.U. and becomes, say, 1.7 A.U. The same data yield 1.1 A.U. as the oxygen radius, so that the oxygen to oxygen distance may be taken to be 2.2 A.U.

(3) Two of Pauling's (73) coördination rules are applied in order to make a choice between two atomic arrangements which give the same calculated relative intensities and which agree with the intensities that are observed.

During the analysis, Thewlis kept in mind the possibility that the oxygen atoms already present in magnetite may have been rearranged while four more oxygen atoms were added during oxidation. He found no evidence of any such rearrangement.

The final result was that the structure of  $\gamma$ -ferric oxide is the same as that of magnetite with oxygen atoms added to the unit cell in the following positions:

0.125, 0.125, 0.125	0.625, 0.375, 0.875
0.875, 0.625, 0.375	0.375, 0.875, 0.625

The positions of the  $\text{Fe}^{++}$ ,  $\text{Fe}^{+++}$ , and  $\text{O}^{--}$  ions for magnetite are completely listed in Thewlis' paper.

The cube edge is 8.40 A.U. long as in magnetite and the space group of  $\gamma$ -ferric oxide is  $O^6$  or  $T^4$ . Thewlis remarks that another structure is possible belonging to space group  $O^7$ , since  $O^7$  is enantiomorphous with  $O^6$ . Their structures are mirror images of each other and are indistinguishable by x-ray methods. The structure diagram for  $\gamma$ -ferric oxide is shown in figure 7.

Thewlis noted that some of the iron atoms have the unusual coördination number of five. Of this he has the following to say: —“A striking feature of the structure is that certain iron atoms (those occupying the  $\text{Fe}^{++}$  positions of  $\text{Fe}_3\text{O}_4$ ) are placed between five oxygen atoms. An arrangement of this kind is most unusual and has to the author's knowledge only once previously been noted, namely, in the case of andalusite (92), where an aluminum atom was found to occur between five oxygen atoms. The structure may with regard to its coördination properties, be regarded as intermediate between  $\text{Fe}_3\text{O}_4$  and  $\alpha\text{-Fe}_2\text{O}_3$ . The fivefold oxygen groups of  $\gamma\text{-Fe}_2\text{O}_3$  can be considered as intermediate between the fourfold groups of  $\text{Fe}_3\text{O}_4$  and the sixfold groups of  $\alpha\text{-Fe}_2\text{O}_3$ .”

A brief résumé of the checkered history of  $\gamma$ -oxide may be of interest. Nearly a hundred years ago, 1838, von Kobell (100) recorded the fact that

the variety of goethite, now called lepidocrocite, becomes magnetic on heating, and in 1848 Plücker (74) noted the same behavior in an artificial ferric hydrate (10). These early observations had been entirely forgotten. The later work of Robbins (77), Malaguti (65) and Liversidge (62) was not generally accepted. This opinion of Dana (24), expressed in 1867, is typical of the skeptical attitude: —“Now, if a sesquioxyd occurs in *isometric* crystals as supposed to be true of  $\text{Fe}_2\text{O}_3$  (but reasonably doubted) that sesquioxyd is not  $\text{Fe}_2\text{O}_3$  but may be . . . .” The presence of ferrous iron was suspected. In 1909, Hilpert (42) proved that ferrous iron need not be present in a ferromagnetic iron oxide. In 1917, Sosman (89) wrote at length of the theoretical interest and importance of ferromagnetic ferric oxide, but it was not until 1925 that much attention was paid to this interesting substance.

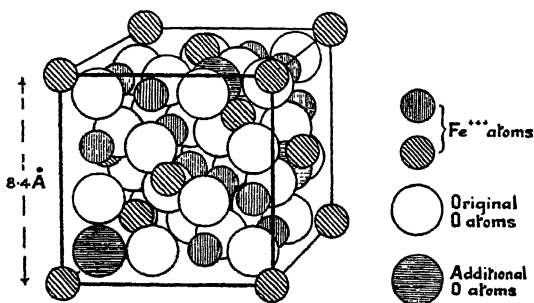


FIG. 7. UNIT CUBE OF  $\gamma\text{-Fe}_2\text{O}_3$   
(Thewlis)

In February, 1925, Welo and Baudisch (103) reported that they had found the x-ray diffraction patterns of oxidized magnetite and of the original magnetite to be identical and that ferromagnetism disappears during the transformation to the usual hematite structure that occurs at a higher temperature. In August of that year appeared their complete paper, a part of which has already been quoted in this review, and the paper of Sosman and Posnjak (90), who had obtained the same experimental results in 1922.

In this same paper Sosman and Posnjak announced that the dehydration of the mineral lepidocrocite,  $\text{Fe}_2\text{O}_3 \cdot \text{H}_2\text{O}$ , yields a ferromagnetic ferric oxide, while similar dehydration of goethite, which is also  $\text{Fe}_2\text{O}_3 \cdot \text{H}_2\text{O}$ , yields only paramagnetic ferric oxide. As has already been mentioned, this alternative path to  $\gamma$ -ferric oxide was known long ago and had long been forgotten. Instead of oxidizing magnetite we may obtain  $\gamma$ -ferric oxide by way of lepidocrocite or its artificial counterpart,  $\gamma$ -ferric oxide hydrate. It is this path, opened up by Sosman and Posnjak, that will now be considered.

# V. CLASSIFICATION OF FERRIC OXIDE HYDRATES: THE YELLOW AND THE RED SERIES

Synthetic  $\gamma$ -ferric oxide hydrate is an orange-yellow hydrate. This statement is much more than a description of its color. The statement identifies  $\gamma$ -ferric oxide hydrate as a member of one of the two great classes of ferric oxide hydrates now recognized, the yellow series and the red series.

TABLE 1

*Comparison of the properties of yellow and red ferric oxide hydrates*

	YELLOW FERRIC OXIDE HYDRATE	RED FERRIC OXIDE HYDRATE
General method of preparation..	Oxidation of $\text{Fe}(\text{OH})_3$ , $\text{FeCO}_3$ , or $\text{Fe}_3\text{O}_4 \cdot$ $x\text{H}_2\text{O}$ with $\text{O}_2$ , $\text{H}_2\text{O}_2$ , $\text{KClO}_3$ , $\text{KIO}_3$ , etc., in presence of water	Precipitation from a ferric salt solution with an alkali
State of hydration.....	Monohydrate	No definite hydrate
Dehydration isobar (47, 86)....	Discontinuous	Continuous
Structure (19, 47, 86).....	Orthorhombic, 2 modi- fications, $\alpha\text{-Fe}_2\text{O}_3 \cdot$ $\text{H}_2\text{O}$ and $\gamma\text{-Fe}_2\text{O}_3 \cdot \text{H}_2\text{O}$	Amorphous
Solubility in acids and ferric chloride solutions.....	Slightly	Easily
Pyrophoricity.....	No	Yes
Isoelectric point (53).....	pH = 5.2	pH = 7.7
Adsorption (37, 91, 101).....	Small	Large
Catalytic activity (decomposi- tion of $\text{H}_2\text{O}_2$ ) (51).....	Low	High
Spontaneous change in catalytic activity (decomposition of $\text{H}_2\text{O}_2$ ) (51).....	Slow	Rapid

They were named in 1879 by Tommasi (94, 96), who was apparently unaware that Muck (70), twelve years earlier, had also observed that two different methods of preparation led to hydrates of identical composition but with different sets of properties. The two general methods of preparation and comparisons of some of their properties are shown in table 1. In this



table, the properties for which no references are indicated are those noted and listed by Tommasi himself.

It will be noticed that the two different modes of preparation are closely analogous to the two ways of forming magnetite that we have already discussed. The reader is referred to that part of this review for a qualitative discussion as to why the yellow hydrates are crystalline substances of definite composition while the red hydrates are amorphous.

Of course the actual line of demarcation between yellow and red hydrates need not be as sharp as indicated in table 1. It is well-known that there is a continuous gradation in color. Sometimes crystal structure may not be detected in a yellow hydrate with x-rays, and red hydrates may contain difficultly soluble yellow hydrates. In all such border-line cases it is found that the fundamental distinction between the two modes of preparation, slow reactions for yellow hydrates and extremely rapid precipitation for the red, has not been maintained. Another reason for the overlapping of properties into the yellow that should be distinctive for the red is that red hydrates convert spontaneously (age) into crystalline  $\alpha$ -ferric oxide hydrate. The several paths of aging are recorded in our diagram, figure 1, where the conventional name of ferric hydroxide is used, and will be considered in some detail later on.

The many special ways of preparing ferric hydrates, especially the yellow ones, will not be considered here, where we are more interested in systematic classification and general relationships. However, because  $\alpha$ -ferric oxide hydrate occurs so rarely,  $\alpha$ -ferric oxide hydrate being the usual product, and because the developments are so recent, we shall consider the chemical conditions necessary for its appearance.

#### VI. PREPARATION AND SOME PROPERTIES OF $\gamma$ -FERRIC OXIDE HYDRATE

Böhm (19) seems to have been the first to recognize one way of preparing  $\gamma$ -ferric oxide hydrate. He mentions casually that the hydrate obtained by the action of oxygen on  $\text{Fe}_2\text{S}_3$  (78) yields the lepidocrocite pattern. Next, Albrecht (4) showed that oxidation of a ferrous salt in dilute solution with an iodate in the presence of sodium thiosulfate (38) yields the  $\gamma$ -modification. Several newer methods have been described by Baudisch (11) and by Baudisch and Albrecht (12). The discussion in the latter paper as to the favorable conditions is as follows:—

“One of us has shown in earlier work (8, 14) that in catalytic oxidations with the aid of iron salts, the oxygen used for the oxidation becomes temporarily activated, becoming coordinatively linked in an intermediate iron complex. Such “active” iron complexes could be identified (9). This supposition contradicts two other theories. Manchot, as the leading representative of one, stipulates the formation of an intermediate iron peroxide, while Wieland, as the supporter of the other, sup-

poses the formation of intermediary hydrogen peroxide. An exact proof for the existence of either the one or the other is lacking at the present time (2), while the facts can be just as well explained by the aid of Baudisch's oxygen-complex theory, with the further advantage that complexes of which oxygen is a member are known with iron as well as with cobalt. If such a complex decomposes on further oxidation, one should expect that ferric oxide hydrate of another structure would be formed in this decomposition, as when, in the direct oxidation of a ferrous salt, the ferric salt formed decomposes by hydrolysis into ferric oxide hydrate and acid.

"This theory explains the following experiment very well. If ferrous chloride solution is oxidized at ordinary temperatures (about 18°), whether quickly or slowly, only  $\alpha$ -hydrate is formed. If, however, the oxidation is carried out *slowly* at a somewhat higher temperature (40–50°) while the solution is being stirred slowly (a saturated solution diluted to one-half), a considerable percentage of  $\gamma$ -hydrate is formed. At temperatures from 90–100°  $\alpha$ -hydrate is obtained with slow as well as with rapid oxidation. The self-complex formation of ferrous chloride solution at higher temperatures has been shown as probable by other investigators (1), and Chandra (21) has succeeded in isolating polynuclear ferrous-ferric complexes as their alkali salts by the partial oxidation of ferrous salt solutions.

"An attempt was then made to obtain pure  $\gamma$ -oxide hydrate by oxidation following the addition of suitable complex-forming compounds to ferrous salt solutions. It has been shown that nitrogen-containing compounds are especially suitable, above all pyridine. As can be seen from the experimental part, it was possible to obtain pure  $\gamma$ -oxide hydrate by the addition of pyridine to ferrous chloride solutions. It has been shown, moreover, that not only the addition of a suitable complex-forming compound, but also the nature of the anions in the salt solution, is of importance for the formation of  $\gamma$ -oxide hydrate, i.e., it was impossible to obtain  $\gamma$ -oxide hydrate from ferrous sulfate or ferrous fluoride solutions. A  $\gamma$ -oxide formation from ferric nitrate solutions was not expected and not found, as the oxidizing power of nitric acid is too great for the formation of any kind of intermediate products with the desired properties.

"In further preparations of  $\gamma$ -oxide hydrate, the addition of sodium azide to the ferrous chloride solutions was found to be of great advantage. The question as to whether this addition of sodium azide to the ferrous salt solutions forms a complex cannot be definitely answered from our experimental material. The red color of the solution points to the formation of an iron azide (63). In oxidations at a higher temperature hydrazoic acid, which can be recognized by its odor, is given off abundantly. It is interesting to note that the formation of  $\gamma$ -hydrate does not commence until the solution is already partly oxidized, and the originally alkaline solution (pH 8 to 10) has become weakly acid. The formation of  $\gamma$ -hydrate first starts from a pH of 6.5 to *ca.* 2.0. It seems then that the formation of  $\gamma$ -hydrate is dependent also on the hydrogen-ion concentration, for it is always found that  $\gamma$ -hydrate is formed only in weakly acidic solutions (pH 2–6.5). This observation is in agreement with that of Krause (58), who is of the opinion that his 'metaferric hydroxide' (eisenige Säure) with the isoelectric point pH 5.2 is identical with the  $\gamma$ -hydrate."

We have only qualitative data as to the relative stability of  $\alpha$ - and  $\gamma$ -ferric oxide hydrate. Albrecht (4) states that  $\gamma$ -ferric oxide hydrate dried for 72 hours at 100°C. had a water content of 6.7 per cent, which is 4.5 per cent less than the theoretical percentage, 11.2, in a monohydrate.  $\alpha$ -Ferric oxide hydrate remained constant in weight at 100°C. with 11.2 per

cent of water and first began to lose it at  $150^{\circ}\text{C}$ .  $\gamma$ -Ferric oxide hydrate is thus less stable than  $\alpha$ -ferric oxide hydrate. In the corresponding minerals, lepidocrocite and goethite, it is also found that the  $\gamma$ -modification (lepidocrocite) is less stable than the  $\alpha$ -form, but the dehydration temperatures are much higher than in the artificial hydrates. In lepidocrocite it is close to  $240^{\circ}\text{C}$ . and in goethite close to  $275^{\circ}\text{C}$ . It must be remembered, however, that the values obtained for dehydration temperatures are much affected by particle size and rate of temperature increase, so that these values are, very likely, too high.

An unpublished experiment of Baudisch shows that  $\gamma$ -ferric oxide hydrate is more easily dehydrated and reduced to form magnetite than is  $\alpha$ -ferric oxide hydrate. The two hydrates were shaken for six to eight

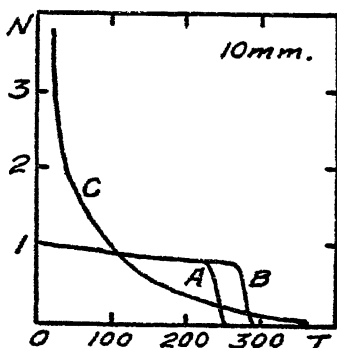


FIG. 8. DEHYDRATION ISOBARS FOR LEPIDOCROCITE, GOETHITE, AND COLLOIDAL FERRIC OXIDE HYDRATE  
(Hüttig and Garside)

hours in glycol that was kept at  $120$ – $125^{\circ}\text{C}$ . A black magnetite was formed from  $\gamma$ -ferric oxide hydrate.  $\alpha$ -Ferric oxide hydrate did not become ferromagnetic and it retained its yellow-brown color, showing that it was not dehydrated either.

Dehydration curves for the natural minerals are shown in figure 8 (44, 47). Curve A is for lepidocrocite,  $\gamma$ -ferric oxide hydrate, and curve B is for goethite,  $\alpha$ -ferric oxide hydrate. In this figure is shown also a curve, C, for an artificial red ferric hydrate for which no discontinuity was expected because no definite hydrates should be present. It is included in the figure to illustrate the fact that entirely different types of dehydration isobars are obtained from true hydrates of the red series. The adsorbed water obeys the laws of osmotic pressure (45, 85).

No such dehydration isobars have yet been taken for artificial  $\gamma$ -ferric oxide hydrate. The evidence for its existence is that it gives the same

x-ray spectrum as lepidocrocite and behaves the same way magnetically, in that it is converted to ferromagnetic  $\gamma$ -ferric oxide on dehydration at low temperatures. Goethite and artificial  $\alpha$ -ferric oxide hydrate do not pass through the ferromagnetic  $\gamma$ -ferric oxide stage but convert directly to  $\alpha$ -ferric oxide.

A proof that artificial  $\gamma$ -ferric oxide hydrate really has one molecule of water has been obtained by Welo and Baudisch (106). A sample made by the pyridine method which had been air-dried contained 16.3 per cent of water. On drying for five months over calcium chloride this was reduced to 15.6 per cent. When dehydrating at  $180^{\circ}\text{C}$ . the water content approached the constant value 4.4 per cent. The difference,  $15.6 - 4.4 = 11.2$  per cent, is the theoretical water content in a monohydrate of ferric oxide. The foregoing data, when viewed in the light of the general form of such dehydration curves as those of Posnjak and Merwin (75), show that the minimum dehydration temperature could not have been much less than the  $180^{\circ}\text{C}$ . that was used.

The crystallographic constants of  $\gamma$ -ferric oxide hydrate as determined by x-ray analysis are, according to Böhm (19):  $a = 3.85$ ;  $b = 12.5$ ;  $c = 3.07$  A.U., giving  $a:b:c = 0.308:1:0.246$ ; the density is 3.97 g. per cubic centimeter. The corresponding values for  $\alpha$ -ferric oxide hydrate are:  $a = 4.60$ ;  $b = 10.01$ ;  $c = 3.04$  A.U.; with  $a:b:c = 0.459:1:0.303$ . Its density is 4.195 g. per cubic centimeter. Virtually the same values were obtained by Goldsztaub (35).

The difficult task of establishing values for the magnetic susceptibilities of  $\gamma$ - and  $\alpha$ -ferric oxide hydrate has been accomplished by Albrecht and Wedekind (4, 6). Both hydrates give  $\chi = +42 \times 10^{-6}$  at room temperature. They conform to the Hausknecht (39) rule that the hydrates have higher susceptibilities than their anhydrides. In  $\alpha$ -ferric oxide,  $\chi = +19.7 \times 10^{-6}$ , and it is independent of temperature (84). Both  $\gamma$ - and  $\alpha$ -ferric oxide hydrate, on the contrary, obey the Curie-Weiss law of temperature variation (5).

#### *$\gamma$ -Ferric oxide from $\gamma$ -ferric oxide hydrate by dehydration*

Williams and Thewlis (107) studied the changes that occur in crystal structure when  $\gamma$ -ferric oxide hydrate is heated at a series of increasing temperatures. Their crystals had, initially, average linear dimensions of the order  $10^{-5}$  cm. When crystals of  $\gamma$ -ferric oxide first appeared they were extremely minute with linear dimensions of the order  $10^{-7}$  cm. A similar disintegration has been observed during dehydration of lepidocrocite by Goldsztaub (35). On heating at higher temperatures the crystals grew to final linear dimensions of the order  $10^{-6}$  cm. At still higher temperatures,  $\gamma$ -ferric oxide transformed to  $\alpha$ -ferric oxide, the crystals of which attained final dimensions of  $10^{-4}$  cm.

throughout a much larger quantity of dextrin. The ferric oxide content of the dry dextrin-oxide mixture was 6 per cent. The results of the comparison are shown in table 2 where the susceptibilities and the remanences are for unit mass of ferric oxide. Again it is seen that the susceptibilities are higher when the crystals or particles are small, and that the more finely divided material has the smaller remanence.

It has been suspected for some time (66) that ferromagnetism might be dependent on the state of division of the material within the range of extreme dispersion, and at the present time the problem of the nature, sizes, and properties of the "magnetic domains" is actively studied. Perhaps further study of the finely divided  $\gamma$ -ferric oxide, with linear dimensions of the order  $10^{-7}$  cm., according to Williams and Thewlis, and even less, according to Welo and Baudisch, would yield important data bearing on this problem.

TABLE 2  
*Susceptibilities and remanences in  $\gamma$ -Fe<sub>2</sub>O<sub>3</sub> and derived powders for colloids*

SUBSTANCE	MAXIMUM SUSCEPTIBILITY $\chi$	REMANENCE $\sigma_r$
		<i>c.m.u.</i>
$\gamma$ -Ferric oxide.....	0.124	1.80
Same, from ball mill.....	0.215	0.635
Same, from colloid mill.....	0.207	1.025

We wish at this point to mention and to illustrate the behavior of colloidal  $\gamma$ -oxide solutions in magnetic fields. Without the magnetic field the usual form of the Brownian movement is observed. Figure 11 is an instantaneous microphotograph of such a solution. The average particle size was 0.1 micron. When the magnetic field is applied chains of particles are formed as shown in figure 12. The Brownian movement displays itself in a different way. The chain as a whole remains practically fixed, but the chains bend at various places from time to time and immediately straighten themselves again. It has also been observed that the length of a chain is a function of the strength of the field.

Before bringing together  $\gamma$ -ferric oxide, derived by way of  $\gamma$ -ferric oxide hydrate, and  $\gamma$ -ferric oxide, derived by low temperature oxidation of magnetite, for a joint discussion of their transformation to  $\alpha$ -ferric oxide, we shall mention the fact that  $\gamma$ -ferric oxide hydrate does not necessarily transform to  $\alpha$ -ferric oxide by way of  $\gamma$ -ferric oxide. It may transform directly to  $\alpha$ -ferric oxide if dehydration is conducted in another way. This way is to grind it. Baudisch and Welo (18) put some  $\gamma$ -ferric oxide

hydrate in a ball mill which was operated for three weeks and observed a change in color from orange-yellow to brick-red. When heated at 180–200°C. this red product did not become ferromagnetic, showing that it was



FIG. 11. PHOTOMICROGRAPH OF  $\gamma$ -OXIDE COLLOIDAL SOLUTION WITHOUT MAGNETIC FIELD

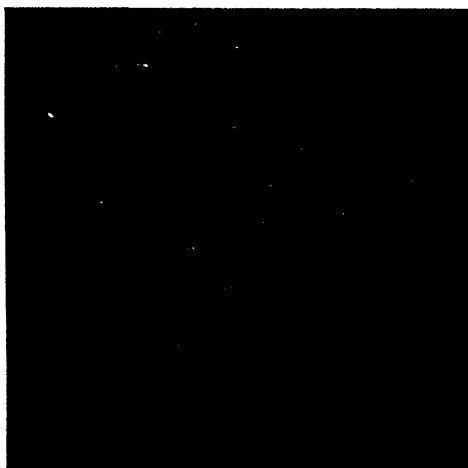


FIG. 12. PHOTOMICROGRAPH OF  $\gamma$ -OXIDE COLLOIDAL SOLUTION IN A MAGNETIC FIELD

no longer  $\gamma$ -ferric oxide hydrate. It might have been either  $\alpha$ -ferric oxide hydrate or  $\alpha$ -ferric oxide. X-ray analysis proved it to be the latter. Ferromagnetism was not detected during the period of grinding so that conversion by way of  $\gamma$ -ferric oxide is excluded.

VII. THE TRANSFORMATION OF  $\gamma\text{-Fe}_2\text{O}_3$  TO  $\alpha\text{-Fe}_2\text{O}_3$ .

Figure 13 shows the typical magnetic behavior of ferromagnetic, cubic  $\gamma$ -ferric oxide as it transforms irreversibly to paramagnetic, rhombohedral  $\alpha$ -ferric oxide when the temperature is raised, or if heated long enough at a sufficiently high temperature (106).

A definite temperature cannot be assigned because the transformation is monotropic and not enantiotropic. Nevertheless, the temperature necessary for a given rate of transformation varies so much from material to material that the cause of the variation assumes a certain interest.

The general experience is that pure materials transform more rapidly at a given temperature. The pure  $\gamma$ -ferric oxide used to obtain the data of figure 13 was studied in a separate experiment. At  $250^\circ\text{C}$ . its permeability fell at the average rate of 0.005 unit per hour during the first 72 hours. Even at  $200^\circ\text{C}$ . a definite decrease in permeability could be detected.

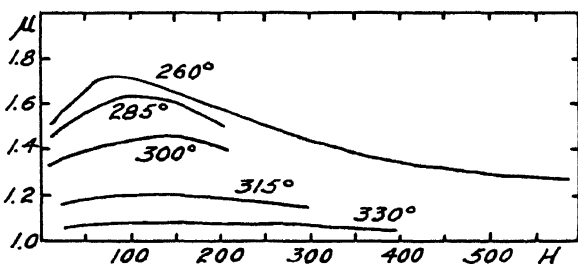


FIG. 13. DISAPPEARANCE OF FERROMAGNETISM ON HEATING  $\gamma\text{-Fe}_2\text{O}_3$

Emmett and Love (28) found that 1.5 per cent of aluminum oxide (catalyst promoter) stabilizes the structure of  $\gamma$ -ferric oxide. The permeability of a certain  $\gamma$ -ferric oxide without aluminum oxide fell to unity in 3 hours when heated at  $450^\circ\text{C}$ . With aluminum oxide in addition the permeability was still 1.10 after heating 18 hours at the same temperature. A  $\gamma$ -ferric oxide from magnetite that had been precipitated with sodium hydroxide instead of ammonia was still more stable. Heating for 18 hours at  $550^\circ\text{C}$ . reduced the permeability to only 1.2.

The same stabilizing effect whenever sodium hydroxide was involved during the preparation of  $\gamma$ -ferric oxide was observed by Huggett (43), and he showed by thermomagnetic analysis that small amounts of sodium ferrite,  $\text{Na}_2\text{O} \cdot \text{Fe}_2\text{O}_3$ , were present. Still another factor which determines the stability was noted by him. The magnetite which he used for the preparation of  $\gamma$ -ferric oxide by low temperature oxidation, was formed by the reduction of  $\alpha$ -ferric oxide with hydrogen in the presence of water vapor. When the reducing temperature was low, for instance,  $300^\circ\text{C}$ ., the

$\gamma$ -ferric oxide subsequently formed by oxidation was much more stable than when the reducing temperature had been high, 500 or 600°C.

These facts in regard to the stability of  $\gamma$ -ferric oxide are neatly combined with a demonstration that the transformation,  $\gamma$ -ferric oxide to  $\alpha$ -ferric oxide, is exothermal in Huggett's thermal analysis curves which are reproduced in figure 14. A silica tube containing  $\gamma$ -ferric oxide was placed in a furnace and a thermocouple inserted in the oxide. Beside this tube was placed another containing chromic oxide, also with a thermocouple. The thermocouples were oppositely connected to a galvanometer. When the transformation occurred, the temperature of  $\gamma$ -ferric oxide increased, and the galvanometer registered the peaks shown in figure 14. In all three cases, A, B, and C,  $\gamma$ -ferric oxide was formed by oxidizing magnetite at 250°C. In case A, magnetite was formed by reduction of  $\alpha$ -ferric oxide

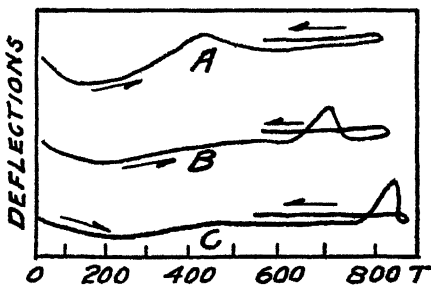


FIG. 14. THERMAL ANALYSES OF  $\gamma$ -Fe<sub>2</sub>O<sub>3</sub>  
(Huggett)

at 500°C. and in case B at 300°C. In both these cases, A and B,  $\alpha$ -ferric oxide had been formed by precipitating ferric nitrate with ammonia and calcining at 750°C. In case C, the original precipitation had been carried out with sodium hydroxide instead of ammonia. The temperature of reduction,  $\alpha$ -ferric oxide to magnetite, was 500°C. as in case A. Huggett also observed that the stability of  $\gamma$ -ferric oxide is influenced by the purity of the ferric nitrate used as starting material in its preparation.

We may note that when the transformation temperature is above the Curie point, the temperature of reversible transformation, it can be directly determined only by a thermal analysis such as in figure 14. For  $\gamma$ -ferric oxide the Curie point is 620°C. Figure 15 illustrates this. The ordinates are proportional to the magnetizations. A and C refer to the same oxides as in figure 14. Oxide A transforms to paramagnetic  $\alpha$ -ferric oxide before the Curie point is reached and is never ferromagnetic when the temperature falls from slightly above 620°C. Oxide C transforms beyond the Curie



point of 620°C. If the transformation temperature has not been reached the magnetization begins to rise as the oxide passes 620°C. on cooling.

It should be noted that in Huggett's curves the magnetizations were observed at varying temperatures. Elsewhere in this review, unless otherwise mentioned, the magnetic properties are described as they appear at room temperature and any reference to temperature means that the

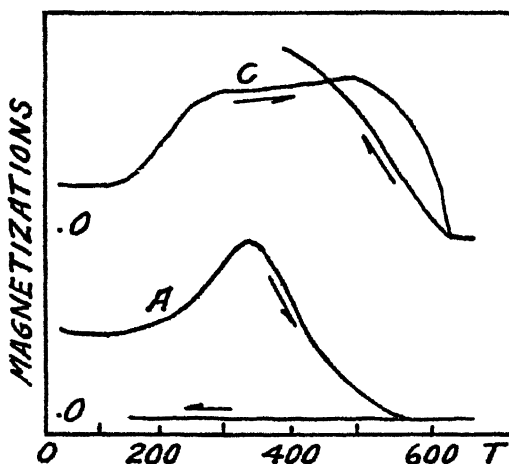


FIG. 15. THERMOMAGNETIC ANALYSES OF  $\gamma$ - $\text{Fe}_2\text{O}_3$   
(Huggett)

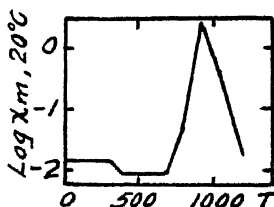


FIG. 16. EFFECT OF PREVIOUS HEATING ON THE MAGNETIC SUSCEPTIBILITY OF  
NONTRONITE  
(Geilmann, Klemm, and Meisel)

magnetic properties were measured *after* heating to the temperature that is indicated.

An analogous and still firmer stabilization in ferromagnetic ferric oxide has been reported by Geilmann, Klemm, and Meisel (31). The natural ferric silicate nontronite,  $[\text{Fe}_2\text{O}_3 \cdot 3\text{SiO}_2 \cdot \text{H}_2\text{O}] \cdot 4\text{H}_2\text{O}$ , was heated at various temperatures and measured magnetically after cooling to 20°C., with results shown in figure 16. The first susceptibility change corresponds to the

elimination of water of constitution. After heating at 900°C. the susceptibility is one hundred times its original value. The x-ray analysis indicated the presence of several intermediate phases at higher temperatures, one of which was probably  $\gamma$ -ferric oxide. At such temperatures as shown,  $\gamma$ -ferric oxide could not exist alone. The lines of  $\alpha$ -ferric oxide first appeared when the temperature had exceeded 1000°C.

Sachse and Haase (81) found  $\gamma$ -ferric oxide to become less magnetic when left to stand at room temperature. During four years the susceptibility fell from  $\chi = 0.250$  to  $\chi = 0.116$ .  $\gamma$ -Ferric oxide aged for this length of time shows an enormously increased rate of transformation at a higher temperature as compared with the fresh substance. Their half-value times, that is, the times of heating necessary to reduce the susceptibilities by one-half, are given in figure 17 for fresh and aged samples as functions of the

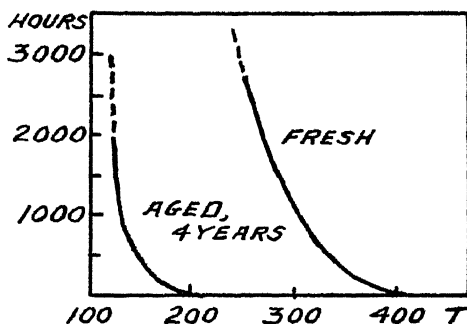


FIG. 17. HALF-VALUE TIMES FOR THE TRANSFORMATIONS OF FRESH AND AGED  $\gamma\text{-Fe}_2\text{O}_3$  TO  $\alpha\text{-Fe}_2\text{O}_3$   
(Sachse and Haase)

temperature. In neither the fresh nor the aged substance did they observe true exponential relations between the susceptibilities and the time of heating at a given temperature. The aged material shows an induction period which is entirely missing in the fresh. Figure 18 shows type forms of curves for fresh and aged  $\gamma$ -ferric oxide.

It was mentioned during the discussion of Huggett's work that the stability of  $\gamma$ -ferric oxide made by oxidizing magnetite depends on the temperature at which the magnetite was formed by reduction of  $\alpha$ -ferric oxide. A similar result was obtained by Luyken and Bierbrauer (64) who experimented with the conversion of low grade limonite ores into  $\gamma$ -ferric oxide as a preliminary step to ore concentration with the magnet. They found that when the reduction of limonite had been carried out at 500°C. less current was required for the operation of the magnetic separator than when the reduction had been carried out at 700°C. They attribute this

difference to the fact that, in the former case, less  $\gamma$ -ferric oxide was transformed to non-magnetic  $\alpha$ -ferric oxide during the subsequent oxidation of the magnetite.

Luyken and Bierbrauer concluded from their experiments that, since a ferromagnetic ferric oxide could be formed from limonite only by way of a preliminary reduction to magnetite, there was no advantage, in this respect, over leaving the ore at the magnetite stage. However, they cite two advantages in reoxidizing magnetite to  $\gamma$ -ferric oxide. The danger is avoided of carrying the reduction too far so as to form weakly magnetic products that do not go into the concentrate and there is, in many cases, a distinct gain in efficiency.

These authors also mention the precautions that are sometimes taken in ore concentration plants employing magnetic separation, to prevent oxida-

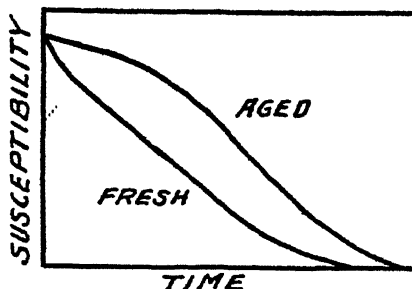


FIG. 18. TYPICAL SUSCEPTIBILITY-TIME RELATIONS FOR THE TRANSFORMATIONS OF FRESH AND AGED  $\gamma$ - $\text{Fe}_2\text{O}_3$  TO  $\alpha$ - $\text{Fe}_2\text{O}_3$   
(Sachse and Haase)

tion of the ore while it is cooling off after reduction. These precautions are taken on the supposition that the ore becomes non-magnetic on oxidation. It is pointed out that the ferromagnetic ferric oxide is formed and that the precautions are quite unnecessary.

#### VIII. VARIOUS FACTS ABOUT $\text{Fe}_3\text{O}_4$ AND $\gamma$ - $\text{Fe}_2\text{O}_3$

##### *Variability in magnetic properties (104, 105)*

Examples have already been given of variable magnetic properties in magnetite prepared in different ways (figure 3). Further examples are given in figure 19. Figure 20 shows correspondingly large variations in  $\gamma$ -ferric oxide. Remanence data for both magnetite and  $\gamma$ -ferric oxide are given in table 3.

The magnetites and  $\gamma$ -ferric oxides used to obtain the curves of figures 19 and 20 and the data of table 3 were prepared as follows:

*Magnetites*.—(1) By precipitation of an equimolar mixture of  $\text{Fe}(\text{SO}_4)_2$  and  $\text{Fe}_2(\text{SO}_4)_3$  with sodium hydroxide. The product was not pure.

(2) By burning iron carbonyl in a limited supply of air. This magnetite was pure and the Fe:O ratio was the theoretical ratio in magnetite.

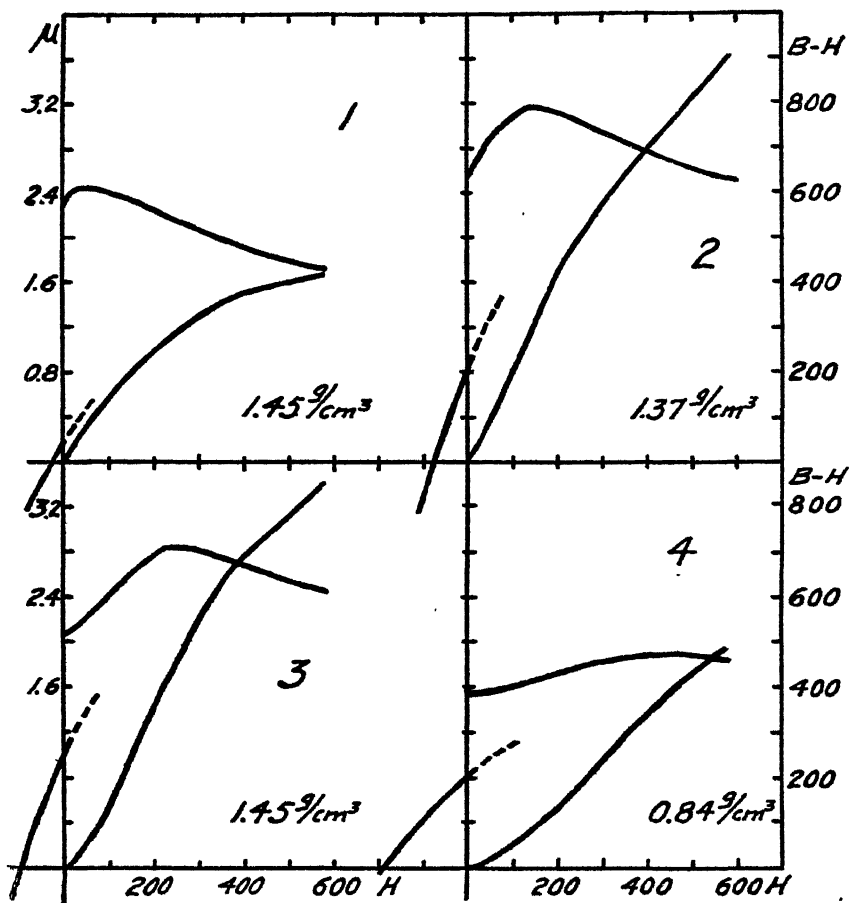


FIG. 19. MAGNETIC PROPERTIES OF MAGNETITES

(3) Magnetite No. 1 was oxidized to  $\gamma$ -ferric oxide in oxygen at  $200^\circ\text{C}$ . and then reduced in molten sodium acetate at  $320^\circ\text{C}$ . for 10 minutes. The product was not pure.

(4)  $\alpha$ -Ferric oxide was made by burning iron carbonyl in a plentiful supply of air and then reduced as in No. 3. The only impurity was carbon from the sodium acetate.

*γ-Ferric Oxides.*—(5) Magnetite No. 1 was heated in oxygen at 200°C. The product was not pure.

(6) Magnetite No. 4 was reoxidized in molten potassium nitrate at 360°C. for ten minutes.

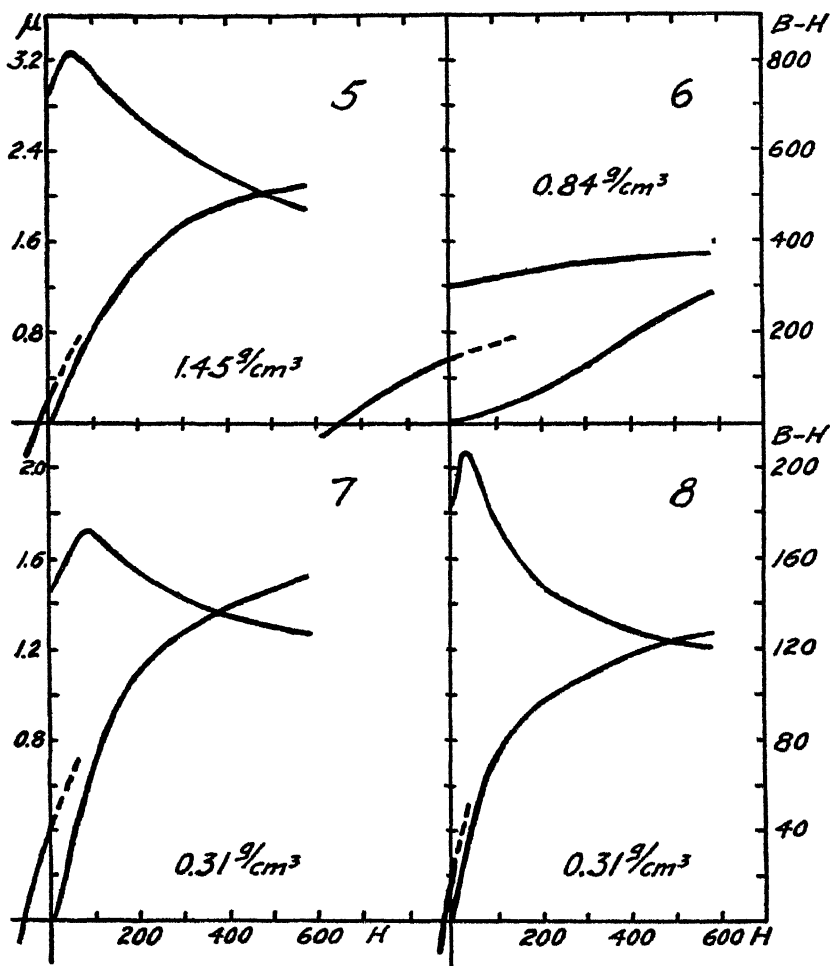


FIG. 20. MAGNETIC PROPERTIES OF FERROMAGNETIC FERRIC OXIDES

(7) By dehydration of  $\gamma$ -ferric oxide hydrate prepared by passing oxygen through a mixture of ferrous chloride and pyridine. The product was very pure.

(8) Same as No. 7 except that ferrous iodide was used instead of ferrous chloride.

From these data and others of the same character Welo and Baudisch (105) have concluded that: (1) There are no major differences in the magnetic properties of magnetite and  $\gamma$ -ferric oxide. A minor difference that always appears, a slightly higher maximum permeability and a shift, to the left, of the permeability maximum in  $\gamma$ -ferric oxide as compared with magnetite, is shown in figure 6. (2) The effect of impurities is a minor one. (3) The magnetic properties are largely determined by crystalline condition, either as to the sizes of the crystals or their perfection, or both. These conditions, in turn, are determined by the method of preparation and subsequent history as to temperature.

TABLE 3  
*Remanences in  $\text{Fe}_3\text{O}_4$  and  $\gamma\text{-Fe}_2\text{O}_3$*

REFERENCE NUMBER*	CLASS OF MATERIAL	DENSITY OF PACKING	REMANENCE $\sigma_r$
		<i>g. per cc.</i>	<i>e.m.u.</i>
1	$\text{Fe}_3\text{O}_4$	1.51	1.97
2	$\text{Fe}_3\text{O}_4$	2.24	15.8
3	$\text{Fe}_3\text{O}_4$	1.58	16.2
4	$\text{Fe}_3\text{O}_4$	1.15	17.0
5	$\text{Fe}_2\text{O}_3$	1.69	1.80
6	$\text{Fe}_2\text{O}_3$	1.57	15.5
7	$\text{Fe}_2\text{O}_3$	—	—
8	$\text{Fe}_2\text{O}_3$	0.69	7.37

\* The reference numbers are those of figures 17 and 18.

#### *Trace of magnetite in $\gamma$ -ferric oxide*

With his thermomagnetic analysis, Huggett (43) found minute traces of magnetite in all of his samples of  $\gamma$ -ferric oxide, although, in agreement with all observers, he found no ferrous iron with the most sensitive chemical tests. This detection depends on the fact that there is a discontinuity in the thermomagnetic curve at  $570^\circ\text{C}$ ., the Curie point of magnetite. A typical case is shown in figure 21. The discontinuity at  $675^\circ\text{C}$ . is characteristic of  $\alpha$ -ferric oxide. Huggett considers the magnetite to be in solution in  $\gamma$ -ferric oxide and remarks that its presence cannot be detected until after the transformation to  $\alpha$ -ferric oxide. The small discontinuity is not visible when superposed on a thermomagnetic curve for  $\gamma$ -ferric oxide, which must be taken at low sensitivity. It would be interesting to see if magnetite is present in  $\gamma$ -ferric oxide made by dehydration of  $\gamma$ -ferric oxide hydrate. Huggett's samples were made by oxidizing magnetite.

$\text{Fe}_3\text{O}_4$ ,  $\gamma\text{-Fe}_2\text{O}_3$ , and the Heisenberg theory of ferromagnetism

According to the Heisenberg theory (41) a condition necessary for the appearance of ferromagnetism in a crystal is that each atom must have at least eight nearest neighbors. Sachse (80) applied this test to magnetite,  $\gamma$ -ferric oxide, and  $\alpha$ -ferric oxide, and found that the condition is satisfied by magnetite and  $\gamma$ -ferric oxide, but is not satisfied by  $\alpha$ -ferric oxide. The same result was obtained by Thewlis (93).

*Photoelectric effect in  $\gamma$ -ferric oxide*

Schwarz (83) has found the photoelectric current from  $\gamma$ -ferric oxide made by oxidizing Lefort's magnetite to be larger than from  $\alpha$ -ferric oxide and from iron powder. He employed the full radiation from the quartz

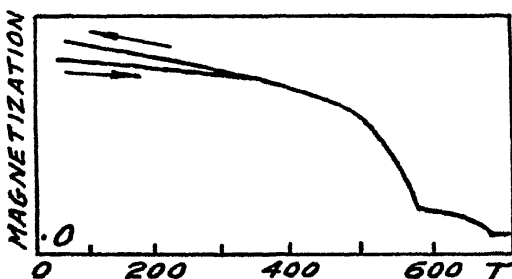


FIG. 21. THERMOMAGNETIC ANALYSIS SHOWING PRESENCE OF  $\text{Fe}_3\text{O}_4$  IN  $\gamma\text{-Fe}_2\text{O}_3$  (Huggett)

mercury arc and observed the following relative values for the photoelectric currents: pure  $\gamma$ -ferric oxide, 6.5; technical  $\gamma$ -ferric oxide, 4.3;  $\alpha$ -ferric oxide made by burning iron carbonyl, 2.5; iron powder formed by dissociating iron carbonyl, 1.6; and ferrum reductum, 1.0.

*Technical  $\gamma$ -ferric oxide or "Siderac"*

A  $\gamma$ -ferric oxide is manufactured and marketed by the Auer Gesellschaft under the name of "Siderac." It is a Lefort magnetite (61, 104) which has been oxidized at about 200°C. until free of ferrous iron.

IX. COLLOIDAL OXIDE HYDRATES

The classical experiments of van Bemmelen (98, 99) have firmly established the fact that there need be no definite stoichiometric proportions between the anhydride and water in colloidal oxide hydrates of multivalent metals, and the capillary theory of Zsigmondy (112) has provided a satisfac-

tory theoretical interpretation. The question as to the chemical constitution of the capillary system's solid framework was, however, left open. It could be either the anhydride or one of several stoichiometric hydrates. The view that the framework is anhydrous is supported by the fact that all of the water in many colloidal oxide hydrates can be reversibly replaced by organic liquids such as alcohol or benzene. In many other colloidal oxide hydrates, however, not all of the water may be so replaced. Several moles of water may remain. There is also the difficulty that many natural minerals corresponding to the colloids whose water is entirely replaceable by organic liquids often appear as definite stoichiometric hydrates.

Foote and Saxton (29) attempted to distinguish water that might be stoichiometrically bound in the framework from all of the other water by measuring the expansion of the colloid on freezing. It was assumed that the water of constitution would take no part in the expansion and that all of the other water would expand as if it were free. It was found, indeed, that the total expansion was less than that corresponding to an equal amount of free water, but that the calculated quantity of combined water did not, in general, correspond to any definite hydrate.

The problem of the chemical nature of the capillary framework assumed a practical as well as theoretical importance when Willstätter and his students (108, 109, 110, 111) began to use colloidal oxide hydrates of aluminum and tin for the selective adsorption and isolation of enzymes. The selectivity depended on the method of preparation and on the age, and it was considered that the selectivity indicated the existence of a number of hydrates in the framework, each having definite adsorptive properties. The well-known method of acetone drying was then developed in an effort to displace the water of adsorption and the water in the capillaries and thus to identify the various stoichiometric hydrates. In this Willstätter was no more successful than Foote and Saxton. In general, stoichiometric relations were not obtained. For example, the percentages of water that remained in a series of ferric hydrogels after acetone drying were 29, 30, 32, and 28 per cent instead of the 33.7 per cent required in a ferric hydrate of the type  $\text{Fe}(\text{OH})_3$ .

Willstätter proposed a new theory. It was assumed that well-defined hydrates are present in the fresh colloids but that they are unstable and changeable, especially in the presence of the capillary water. They either react with each other in the sense of salt formation or else they undergo inter- and intra-molecular dehydration so as to form a complex of various hydrates. Neither the paths of spontaneous dehydration nor the final stages are unique, so that the final product is a very complicated mixture. In such a mixture, stoichiometric relations cannot be observed. The dehydration isobars would closely approach continuous curves. The



chemical changes that occur with time, and especially the changes in adsorptive power towards enzymes, were referred to by Willstätter as being in agreement with the view that numerous new hydrates are formed on aging.

Simon (85) has pointed out that all of these phenomena are equally well explained by the older views as to what occurs when a substance ages and mentions, particularly, the process of aggregation. However, Simon's principal argument against Willstätter's view is that it is possible, by "starke Eingriffe," to form definite hydrates from colloids, which will have crystalline structure and which will give discontinuous dehydration isobars. According to the Willstätter theory we should have, instead,

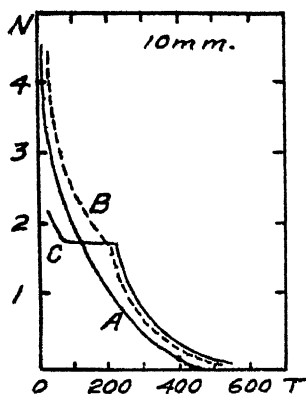


FIG. 22. DEHYDRATION ISOBARS FOR ANTIMONY PENTOXIDE HYDRATE  
(Simon and Thaler)

strong dehydration and should obtain a complicated mixture which would give a continuous curve of dehydration.

This point Simon illustrated by his own and Thaler's (87) work on colloidal antimony pentoxide hydrate. Their dehydration isobars are shown in figure 22. Curve A is for a colloid that had aged for a few months. Another portion of the same sample was studied after aging for about a year and gave curve B. There is observed not only a firmer binding of the water but also an indication of a hydrate with 1.66 moles of water per mole of antimony pentoxide. Further aging for another year brought about little change, so another sample was prepared in the same way and heated with water in a sealed tube for several days at 300°C. This material gave curve C and confirms the supposition that aging of a colloid consists, in part, of the fixing of the water to form a definite hydrate. In this case the hydrate is  $3\text{Sb}_2\text{O}_5 \cdot 5\text{H}_2\text{O}$ . Similar results as to the firmer fixation of water

have been obtained by Simon and Schmidt (86) in their study of colloidal ferric hydroxide, with the important additional result that this fixation of water is accompanied by crystallization to such an extent that it can easily be detected by x-ray methods.

We may sum up the immediately preceding discussion as follows: That in some and most probably in all of the colloidal metal oxide hydrates, the solid framework of the capillary system consists of the anhydride, and there is no evidence that it ever consists of a definite stoichiometric hydrate; and that the result of spontaneous aging is to fix a part of the water more firmly and to form definite hydrates of crystalline structure.

#### X. THE AGING PATHS FROM COLLOIDAL FERRIC OXIDE HYDRATES

Because of the relative free energies in amorphous and crystalline systems it is self-evident that the direction of aging must always be from and never towards the colloidal state and we need only to consider the various paths, the intermediate phases that may appear, and the final stage that is reached. We shall consider, first, the views to which Hüttig and Zörner (52) and others of the Hüttig school have been led.

A slow elimination of water to form hydrohematite is one of the possible paths, since many ferric oxide hydrate colloids that had been aged for a period of years showed all of its characteristics. Hüttig and Zörner cite Ruff's (79) observations on a 6-year old hydrate and describe their own comparative studies on a fresh hydrate, a hydrate that had been prepared in the same way twenty-nine years earlier, and a natural hydrohematite. The dehydration isobars are shown in figure 23 where curve A is for the fresh colloid, curve B is for the 29-year old hydrate, and curve C is for the natural mineral. There is a general similarity in isobars B and C, although the amount of combined water in the aged artificial hydrate is somewhat higher than in the natural. On comparing curves A and B we see that the aging consists, not in a general loosening of all of the water but, rather, in nearly complete elimination of water above 0.8 mole and an even firmer binding of this remainder, since curve B lies above curve A at the higher temperatures of dehydration. Both the 29-year old artificial hydrate and the natural hydrohematite gave the x-ray spectrum of  $\alpha$ -ferric oxide with no determinable departures from the lattice dimensions of true  $\alpha$ -ferric oxide. According to Kurnakow and Rode (60) the combination with water is zeolithic in character, and they regard the water as being in solution in hematite. The water content in hydrohematite is variable and it seems, from the results of both Kurnakow and Rode and Hüttig and Zörner, that the upper limit is at about 0.8 mole per mole of ferric oxide.

Very interesting subsidiary evidence that hydrohematite is one of the intermediate steps in the aging of colloidal ferric hydroxide is found in the magnetic studies of Hüttig and Kittel (48). They consider  $\chi = +116 \times 10^{-6}$  to be the most probable susceptibility value for fresh colloidal ferric hydroxide, and show that it is paramagnetic. Their criterion for paramagnetism is that the susceptibility value shall be independent of the field strength. If the susceptibility decreases with increase of the field, the substance is considered to be ferromagnetic. They then proceeded to show that colloidal ferric hydroxide, on aging, not only assumed higher susceptibilities but that the values became more and more dependent on the field strengths that were used during the measurements. Hydrohematite,

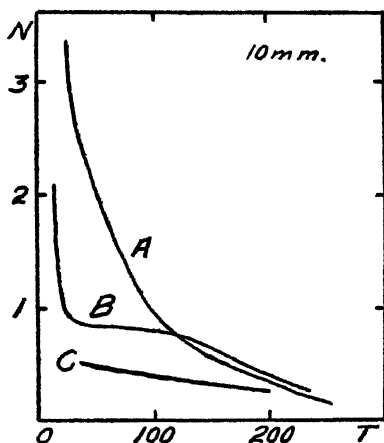


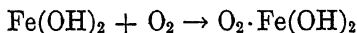
FIG. 23. DEHYDRATION ISOBARS FOR FRESH AND AGED COLLOIDAL FERRIC OXIDE HYDRATES AND OF HYDROHEMATITE  
(Hüttig and Zörner)

therefore, should be ferromagnetic in the sense of the foregoing definition. They found this to be true of both the natural hydrohematite and of the 29-year old artificial hydrate. The susceptibility values for the natural hydrohematite were  $207, 144,$  and  $117 \times 10^{-6}$ , respectively, for magnetizing currents of 0.5, 1.0, and 1.5 amperes. For the 29-year old hydrate the susceptibilities were  $114$  and  $98 \times 10^{-6}$  for magnetizing currents of 0.5 and 1.0 amperes, respectively. Both are therefore ferromagnetic. The smaller susceptibilities for the aged artificial hydrate are probably due to partial conversion to limonite or to goethite. Both of these are paramagnetic with much lower susceptibilities which fall within the range  $31$  to  $44 \times 10^{-6}$  according to Hüttig and Kittel.

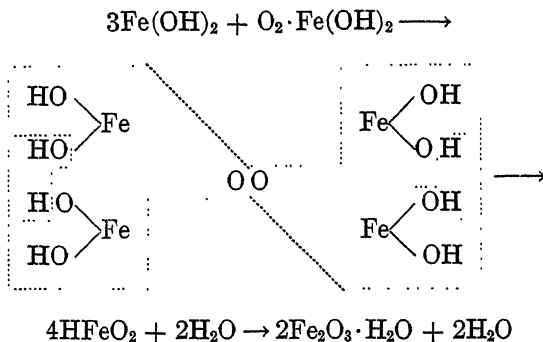
When beginning the discussion of hydrohematite as a product of the aging of colloidal ferric hydroxide it was stated that it was formed by the elimination of water down to about 0.8 mole. Actually the mechanism is not so direct. Otherwise we would not find hydrohematite with less than 0.8 mole of water and it would be difficult to understand how hydrohematite can change to limonite and eventually to goethite,  $\alpha\text{-Fe}_2\text{O}_3 \cdot \text{H}_2\text{O}$ . The probable true mechanism was first described by Ruff (79). He considered that the primary step in the aging process, under ordinary conditions of temperature and pressure, is the elimination of all of the water to form the anhydride, but that this anhydride, in the course of time, dissolves water (zeolithic binding). The amount of zeolithic water may exceed 1 mole. Whenever zeolithic water exceeds or is not too much below 1 mole, recrystallization takes place and we have a gradual conversion to  $\alpha\text{-Fe}_2\text{O}_3 \cdot \text{H}_2\text{O}$ . Limonite is thus an intermediate step in the aging path from hydrohematite to goethite. Hüttig and Zörner (52) consider this to be the only possible path and that hydrohematite cannot convert directly to  $\alpha\text{-Fe}_2\text{O}_3 \cdot \text{H}_2\text{O}$ .

The chief experimental basis for the belief that colloidal ferric hydroxide ages to limonite (goethite crystal structure with zeolithic water) is an old observation of Schiff (82), who found that he could not reduce the water content of a 15-year old hydrate below 11 per cent by heating in air at  $100^\circ\text{C}$ . Eleven per cent of water corresponds very closely to 1 mole of water per mole of ferric oxide. The other evidence, which we owe to Doelter (27), is that colloidal ferric oxide hydrate, like many other colloids, was partly crystallized after having been shaken in water; in this particular case, for 88 days. It should be noted, however, that Doelter's colloid was made by oxidation of ferrous hydroxide, so that his results might not be valid for a colloid made by precipitation from a ferric salt.

Hüttig and Zörner (52) note that limonite is also obtained by the oxidation of ferrous hydroxide, but in a pure or unmixed form only when molecular oxygen is used as the oxidizing agent. Other oxidizing agents such as hydrogen peroxide, in which the ultimate oxidizing medium is atomic oxygen, lead at the most to only a partial formation of limonite. They also note that this oxidation to limonite proceeds best with fresh, highly dispersed, ferrous hydroxide. In their explanation they assign an orienting as well as an oxidizing rôle to the oxygen molecule. It is assumed that the primary process is the formation of a peroxide



and that this peroxide combines with three unoxidized ferrous hydroxide groups according to the scheme:



The structure represented by the second stage of this scheme is regarded as a sort of seed or nucleus which initiates crystallization to goethite,  $\alpha\text{-Fe}_2\text{O}_3 \cdot \text{H}_2\text{O}$ .

We cannot, as we have seen, distinguish limonite from goethite by magnetic measurements. They can be distinguished by their x-ray spectra and the dehydration isobars. While the positions of the lines in the limonite x-ray spectrum are the same as those of goethite,  $\alpha\text{-Fe}_2\text{O}_3 \cdot \text{H}_2\text{O}$ , they are, regardless of source or age, more diffuse and less intense. This is attributed by Hüttig and Zörner to high dispersion and incomplete ordering as long as zeolithic water is present. This excess water must be eliminated before true goethite crystals can be formed. Typical isobars for limonite are shown in figure 24 along with one for goethite. Curve A is for a natural mineral of the limonite class and curve B is for a natural goethite. Limonite does not dissociate into two phases until the composition of goethite is reached with 1 mole of water. Between 1 mole of water and about 0.4 mole the two phases are  $\alpha$ -ferric oxide hydrate and hydrohematite. When the water content is less than 0.4 mole, there are also two phases, hydrohematite and hematite itself. Hüttig and Zörner note that in the range of the two phases,  $\alpha$ -ferric oxide hydrate and hydrohematite, the dissociation pressure of 10 mm. was reached at very close to  $177^\circ\text{C}$ . in all of the limonites that were studied. They consider that the conditions of thermodynamic equilibrium were more nearly fulfilled at this pressure in the limonites than in goethite. When fully developed crystals are present, as in goethite, dissociation is retarded until a higher temperature is reached. The corresponding temperature for the dissociation of  $\alpha$ -ferric oxide hydrate to hydrohematite is about  $100^\circ\text{C}$ . higher in goethite.

In figure 24, curve C is for a fresh hydrate made by the oxidation of ferrous hydroxide according to a procedure followed by Tommasi (95), and curve D is for a hydrate prepared in the same way twenty-nine years earlier. They are of particular interest. We may recall that since these

two hydrates were formed by the oxidation of ferrous hydroxide, they belong to the yellow series, whereas hydrates precipitated from ferric salts are classed as of the red series. We have seen, while comparing the isobars A and B in figure 23, that a red hydrate changes greatly during twenty-nine years. The yellow hydrate, on the other hand, has scarcely changed at all during the same time. It was crystalline and the free energy content was near the minimum from the start. We may infer that it would require an extremely long time, under ordinary temperature and pressure conditions, for the limonites, or rather, yellow hydrates, to assume the final stable state of goethite. It should hardly be necessary to state that while we have been discussing limonite as one member of the aging path from colloidal ferric hydroxide, we have at the same time been considering the conversion of a hydrate of the red series to one of the yellow.

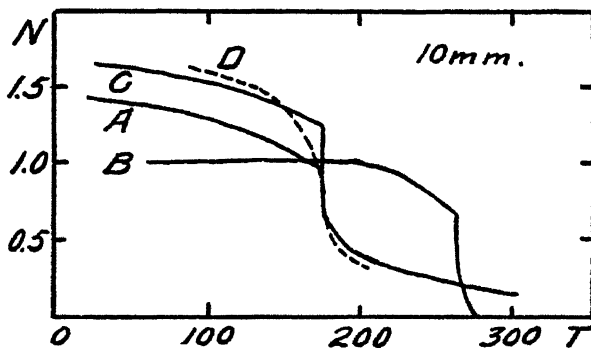


FIG. 24. DEHYDRATION ISOBARS FOR LIMONITE, GOETHITE, AND FRESH AND AGED YELLOW FERRIC OXIDE HYDRATES  
(Hüttig and Zörner)

#### XI. THE SPECIFIC HEATS AND THEIR CHANGES

The specific heats of colloidal ferric oxide hydrates with various amounts of water have been measured by Hüttig and Garside (47). The range of water content covered was from 0.14 to 3.8 moles per mole of ferric oxide. To obtain material with so little water as 0.14 mole, drying at as high a temperature as 200°C. was resorted to. Even after heating to this temperature, the material remained colloidal in the sense that no crystallization had occurred. It was much more readily soluble in hydrochloric acid than crystalline ferric oxide and remained pyrophoric. When they tried to reduce the water content still further, crystallization set in and the material became unsuitable for their purpose, which was to measure the specific heats and to study the chemical and catalytic activities of colloidal (un-crystallized) ferric oxide hydrates with various amounts of water.

The specific heats were measured within two or three days after preparation, so that the data are considered to be for fresh and not for aged material. They are plotted as functions of the water content in figure 25.  $N$  is the number of moles of water per mole of ferric oxide. For  $0.2 < N < 1.0$  the relation is linear. At  $N = 1$ , the specific heat,  $C_p$ , appears to be a mean value between the specific heat of goethite,  $C_p = 0.2040$ , and the calculated value of 0.2458 for an equimolar mixture of ferric oxide and water. These two values are indicated by (g) and (e) in the figure. A discontinuity appears at  $N = 1$ . Beyond  $N = 1$  the curve is steeper as if a part of the water were free. However, the water is not present as an independent phase, since the course of the curve beyond  $N = 1$  is not linear. For  $N < 0.2$  the curve departs from linearity and for

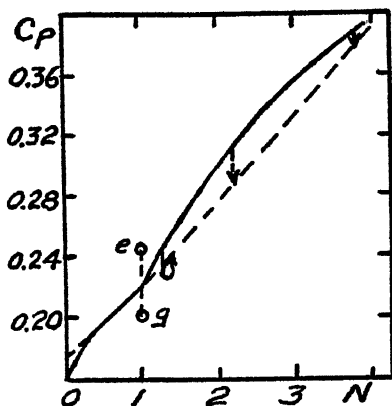


FIG. 25. SPECIFIC HEATS OF COLLOIDAL FERRIC OXIDE HYDRATES  
(Hüttig and Garside)

$N = 0$ ,  $C_p = 0.1608$ , the specific heat of hematite,  $\alpha$ -ferric oxide. Extrapolation of the linear part of the curve to  $N = 0$  yields  $C_p = 0.175$ , which Hüttig and Garside consider to be the specific heat that would be found for ferric oxide if it were possible to completely dehydrate colloidal ferric oxide hydrate without crystallization and without destroying its chemical and catalytic activity.

The vertical arrows of the figure indicate the directions and the magnitudes of the changes in specific heat as the hydrates are allowed to age under ordinary conditions of temperature and pressure. The specific heat of the hydrate with  $N = 3.8$  fell from 0.382 to 0.370 in 80 days. For the hydrate with  $N = 2.1$ , the change was from 0.302 to 0.284 in 78 days. In the hydrate with  $N = 1.2$ , the specific heat decreased from 0.240 to 0.233 during the first 14 days and then increased to the original value during

the following 36 days. A similar but very slight fall and rise was observed in the hydrate with  $N = 0.25$ .

Hüttig and Garside consider a decrease in specific heat to be an indication that the path of aging is towards goethite,  $\alpha$ -ferric oxide hydrate, with firmer binding of the water. An increase in specific heat would then indicate that water is being set free with the formation of crystalline  $\alpha$ -ferric oxide. They note that the former occurs in hydrates rich in water and that the latter is more likely to happen when there is not enough water present to form a monohydrate.

## XII. THE ACTIVITIES OF FERRIC OXIDE HYDRATES AND THEIR CHANGES

The hydrates that were used for specific heat measurements were also used in the study of catalytic and chemical activities. We shall consider, first, these activities as functions of the water contents. As a measure of

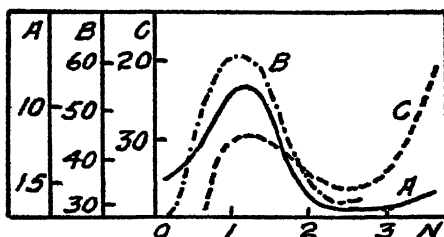


FIG. 26. EFFECT OF WATER CONTENT ON ACTIVITIES OF COLLOIDAL FERRIC OXIDE HYDRATES  
(Hüttig and Garside)

catalytic activity, Hüttig and Garside (47) chose the rate of decomposition of hydrogen peroxide. The time necessary for the evolution of 100 cc. of oxygen with 0.2 g. of ferric oxide was measured, the strength of the hydrogen peroxide solution being 5 cc. of 30 per cent hydrogen peroxide in 50 cc. of water. Chemical reactivity was tested by measuring the volume of hydrogen sulfide that combined with 0.2 g. of ferric oxide in a given time. Finally, the rate of solution in hydrochloric acid was determined by measuring the time required for completely dissolving 0.2 g. of ferric oxide when shaken in 1000 cc. of a half-normal solution. All of these measurements were made within three days from the time the hydrates were prepared, so that the data pertain to "fresh" material.

It is necessary to explain the three scales of ordinates that are used in figure 26 which exhibits the parallelism between catalytic activity and chemical reactivity, and their striking changes with variation in water content. When the catalytic activity is high, less time is required for the



evolution of a given volume of oxygen from hydrogen peroxide. In the figure, therefore, the ordinate scale designated by A reads downward. The unit of time is the minute. The scale of ordinates, B, measures the volume of hydrogen sulfide that combined with ferric oxide. Evidently the scale reads upward in the usual manner. Hüttig and Garside do not state the time adopted as a standard for this reaction. Again, on scale C, on which is represented the rate of solution in hydrochloric acid, we must read downward. The unit is the minute. The letters A, B, and C for the curves, correspond, of course, to those for the ordinates. On the abscissa the unit is the moles of water per mole of ferric oxide.

In regard to the data of figure 26 we shall content ourselves, as did Hüttig and Garside, with a résumé of the facts. In relatively fresh colloidal ferric hydrates, the changes in catalytic activity, chemical reactivity, and rate of solution, in relation to the water contents of the hydrates, run closely parallel. The hydrates with little water have low activities. The activities increase with increase in water content and reach a maximum at about  $N = 1$ . Beyond  $N = 1$  the activities decrease to a minimum between  $N = 2$  and  $N = 3$ . Beyond  $N = 3$  the activity increases again. One would expect to find, in hydrates very rich in water, activities greater than any that were observed in these experiments.

Remarkable to-and-fro changes in the same activities were found in aging ferric oxide hydrates. Hydrates with water contents of 0.73, 1.23, and 3.72 moles were set aside to age, and from time to time portions were examined as to activity as described in the preceding paragraphs. From the data, the activities of a hydrate with a given water content could be plotted as functions of the time of aging. It was found that the oscillatory changes were most marked in the hydrate with the least water. Only the results for  $N = 0.73$  will be shown here (figure 27). The letters A, B, and C have the same significance as in figure 26, but the abscissa measures the time of aging in days.

The interpretation to be given such striking periodic changes in activity is as yet rather vague. This is not surprising, since we are still too much accustomed to the view that the activity function in an aging system is always a continuously decreasing one. The qualitative theory advanced by Hüttig and his associates (46, 47, 50, 51, 52) will be briefly outlined. They consider that a system, already crystalline, is most active when it, or some part of it, is undergoing a change to another crystalline form. They cite as an example the observation of Hedvall and Sjöman (40) that ferric oxide reacts with quartz only at the temperatures of 575 and 950°C. which are, respectively, the temperatures of the transformations  $\beta$ -SiO<sub>2</sub> to  $\alpha$ -SiO<sub>2</sub> and  $\alpha$ -SiO<sub>2</sub> to cristobalite. Now the change in structure generally involves the existence of a highly disperse intermediate phase, so that the total

surface may be very large. Also, the intermediate phase may not be truly crystalline. It may consist of a loose, unordered aggregate in the sense of the Smekal (88) theory of the structure of catalysts. In either case, a change in structure would lead to higher catalytic activity, and it remains to account for the periodic changes in structure in aging ferric oxide hydrate.

As we have seen (figure 1), colloidal ferric oxide hydrate ages along two different paths at first. One is to goethite by way of limonite and the other is to hydrohematite. Actually, both paths are followed by separate portions of the material. According to Hüttig, the part that ages to hydrohematite cannot reach the most stable form, goethite, unless the path is changed to limonite. It is assumed that this change in path can occur spontaneously, and it is at this time of change that we find the high cata-

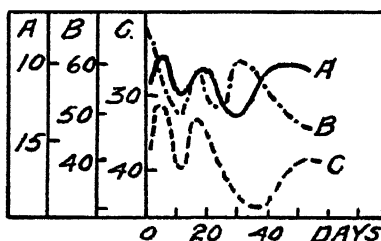


FIG. 27. VARIATION WITH AGE OF THE ACTIVITIES OF COLLOIDAL FERRIC OXIDE HYDRATE  
(Hüttig and Garside)

lytic activity. Of course only a small part of the material may change at a given time. Later on, another portion that has been following the path to hydrohematite may change to limonite and result in another period of high activity. Hüttig and Zörner (51) have observed as many as nine peaks in an activity curve within a period of 160 days.

A curious parallelism between catalytic activity and ferromagnetism has been noted several times. We have already mentioned the fact that hydrohematite has the characteristics of a ferromagnetic substance, while all the other members of the aging series from colloidal ferric oxide hydrate are paramagnetic. The greater catalytic activity of  $\gamma$ -ferric oxide, as compared with  $\alpha$ -ferric oxide, has been observed by Baudisch and Welo (16). Merck and Wedekind (67) found that high catalytic activity and ferromagnetism often appeared simultaneously during the dehydration of cobalt oxide hydrates. It is not yet known whether or not the relation between catalytic activity and ferromagnetism is a causal one.

This parallelism can best be illustrated by the work of Hüttig, Radler,

and Kittel (50) on zinc-chromium spinel. A stoichiometric mixture of zinc oxide and chromic oxide was heated for 6 hours at each of a series of increasing temperatures. After each heating the material was examined as to the catalytic activity towards the dissociation of methanol, the magnetic susceptibility, and the packing volume. The variations in these properties are shown in figure 28. There seems to be no close correspondence between the various maxima and minima. However, it was pointed out by these investigators that the maxima occurred in the temperature range of previous heating, 300 to 500°C. After heating at these temperatures the x-ray spectra showed lines belonging to neither the mixture or

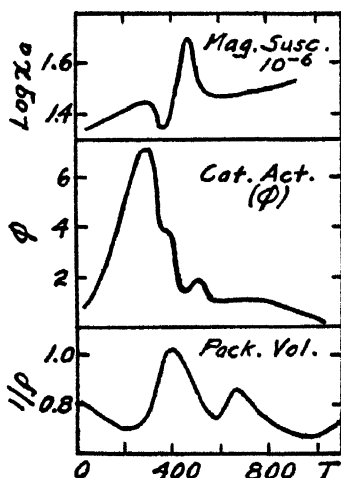


FIG. 28. VARIATIONS OF MAGNETIC SUSCEPTIBILITY, CATALYTIC ACTIVITY, AND PACKING VOLUME DURING THE FORMATION OF ZINC-CHROMIUM SPINEL  
(Hüttig, Radler, and Kittel)

oxides nor the finished spinel. In other words, high activities and high susceptibilities are to be associated in some way with intermediate phases. That these phases are more highly dispersed is indicated by the higher values of the packing volumes.

#### *Activities in yellow ferric oxide hydrates*

The difference between red colloidal and yellow crystalline ferric oxide hydrate as a catalyst in the decomposition of hydrogen peroxide has been illustrated by Hüttig and Zörner (51). An 8-day old red hydrate, made by adding ammonium hydroxide to ferric chloride in the cold, yielded 6 cc. of oxygen per minute, while a hydrate that had been prepared in the same

way twenty-nine years earlier yielded only 0.11 cc. per minute under comparable conditions. A yellow hydrate, formed by adding sodium carbonate to ferrous sulfate in the presence of potassium chlorate, released 0.33 cc. of oxygen per minute when fresh. However, the activity of a yellow hydrate changes extremely slowly with time. One which had been prepared in the same way twenty-eight years earlier still yielded 0.25 cc. of oxygen per minute. As has been stated before, the yellow hydrates are crystalline from the beginning. Perhaps whatever change there was was due to small amounts of amorphous hydrates.

That the temperatures at which yellow hydrates are formed have some effect on the catalytic activity is illustrated by some unpublished measurements of Baudisch. The hydrates were prepared by the action of molecular oxygen on ferrous bicarbonate at temperatures of 0, 19.5, and 37°C. The

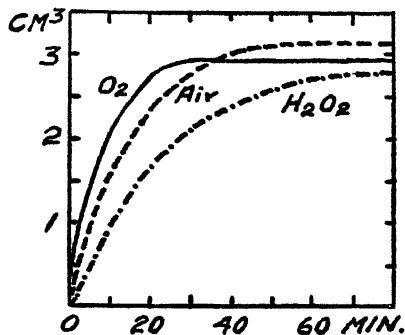


FIG. 29. CATALYTIC ACTIVITIES OF YELLOW FERRIC OXIDE HYDRATE

relative amounts of hydrogen peroxide dissociated under comparable experimental conditions were, respectively, 40.3, 38.8, and 28.0. When the temperature is low the rate of crystallization is slowed up. The crystals may be both smaller and less ordered than when the temperature is high.

Other unpublished work by Baudisch brings out the fact that the method of oxidation used in preparing yellow hydrates affects the activity. They were formed from ferrous bicarbonate solutions in three ways. In one, air was bubbled through the solution; in another, oxygen. In the third, hydrogen peroxide was added to the solution. The total volumes of oxygen that had been liberated from hydrogen peroxide were measured at various times as shown in figure 29. In each case the activity was small and the activities among the hydrates differed at the most by a ratio of about 2.

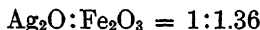
### XIII. THE SILVER FERRITE ANALYSIS AND ITS APPLICATION TO THE STUDY OF FERRIC OXIDE HYDRATES

Very interesting studies on the classification of ferric oxide hydrates and the identification of the products of aging have been conducted by A. Krause and his students during the past two years. Krause's points of view and his methods of attack are closely analogous to those of the organic chemist. Definite molecular structures are assigned to the freshly formed hydrates and changes in them are considered on the bases of polymerization, ring closure, aggregation, oxygen bridging, etc.

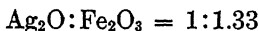
The underlying feature of this method of analysis is the synthesis of silver ferrites and the difference that immediately appears between yellow hydrates and red hydrates, or, as Krause terms them, respectively, meta-hydroxides and ortho-hydroxides. Krause and Pilawski (58) found that the composition of a silver ferrite formed from meta-hydroxide can be represented by the proportion



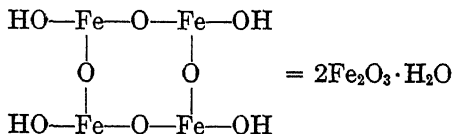
In a silver ferrite formed from ortho-hydroxide, on the other hand, this proportion proved to be



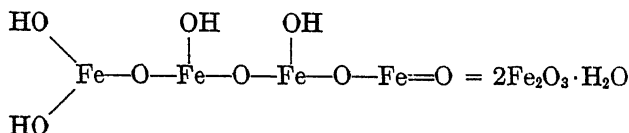
From considerations of a chemical nature, Krause and Pilawski conclude that the true composition of a silver ferrite made from fresh ortho-hydroxide would be



which suggests that meta- and ortho-hydroxides are four-membered molecules. They assign to the meta-hydroxide a ring structure with equivalent hydroxyl groups and in which each hydrogen is replaceable by silver, thus:



To the ortho-hydroxide is assigned a chain structure



in which only three OH groups are equivalent with replaceable hydrogens. The fourth OH group plays a different rôle, and it is this group which imparts to ortho-hydroxide its more basic character as compared to meta-hydroxide. We have already seen (53) that the isoelectric point of the former is at  $\text{pH} = 7.7$ , whereas that of the meta-hydroxide is at  $\text{pH} = 5.2$ .

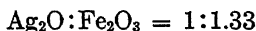
The ring molecule which forms silver ferrite with  $\text{Ag}_2\text{O}:\text{Fe}_2\text{O}_3 = 1:1$  is the  $\gamma$ -ferric oxide hydrate. By the silver analysis it is therefore possible to identify a  $\gamma$ -ferric oxide hydrate which is not sufficiently crystallized to show an x-ray spectrum. This has been done by Krause, Czapska, and Stock (55).  $\alpha$ -Ferric oxide hydrate, either crystalline or amorphous, does not form ferrites with silver. The silver analysis thus provides an additional means of identifying the products of the oxidation of ferrous hydroxide. Professor Krause, in a private communication, informs us that amorphous (no x-ray lines) as well as crystalline  $\gamma$ - and  $\alpha$ -ferric oxide hydrate have been obtained by this reaction. It is for this reason that the oxidation paths, ferrous hydroxide  $\rightarrow$  amorphous ferric acid (amorphe eisenige Säure), and ferrous hydroxide  $\rightarrow$  amorphous  $\alpha$ -hydrate, have been added to the diagram of figure 1. It may be mentioned here that the aging paths and the numerous steps that have been disclosed by Krause and his associates are shown by broken lines in the figure. Paths which were already known are shown with unbroken lines.

As auxiliary aids in the analysis, Krause has used the fact that meta-hydroxide and other ring compounds dissolve extremely slowly in concentrated acetic acid and in 30 per cent nitric acid, while ortho-hydroxide and chain polymers dissolve very easily (53, 59). The fact that meta-hydroxide may be peptized in dilute solutions of either acids or alkalies, while the ortho-variety can be peptized only in acid solutions, is another way of distinguishing between them. Color changes also give valuable indications as to the nature of the changes taking place on aging.

We have seen that silver ferrites formed from meta-hydroxide have the composition



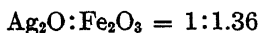
Consequently meta-hydroxide may be given the formula  $(\text{HFeO}_2)_n$ . The ferrite formed from ortho-hydroxide with



has for formula  $\text{Ag}_3\text{H}(\text{FeO}_2)_4$ .

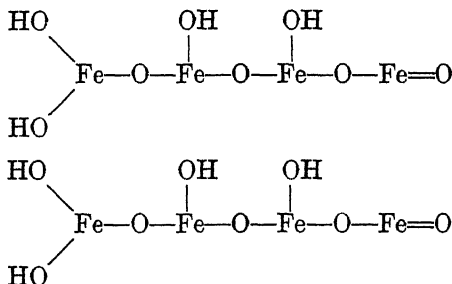
#### XIV. AGING OF ORTHO-HYDROXIDE

The "activity" of ortho-hydroxide and its tendency to age is evident even in the moment of precipitation from a ferric salt with an alkali. Silver ferrites of the composition

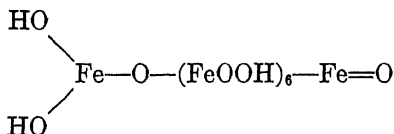


can only be obtained by precipitation in the presence of silver nitrate so that the silver reacts with ferric hydroxide in the nascent state. If silver nitrate is added immediately after precipitation some aging will already have occurred, as evidenced by  $\text{Ag}_2\text{O}:\text{Fe}_2\text{O}_3$  ratios of more than 1:1.36. However, the rate of aging can be controlled by varying the hydroxide concentration. Aging of ortho-hydroxide is most rapid (59) when the alkalinity is between  $N = 0.1$  and  $0.05$ . With  $N = 1$ , the aging is much slower.

According to Krause and Ciokówna (54), the first step in the aging is the polymerization of the four iron atom chain molecules to form longer chains, and it is due to a loosening of the  $\text{Fe}=\text{O}$  double bond. For two ortho-hydroxide molecules this may be illustrated by the union of



to form a molecule of eight iron atoms which may be abbreviated



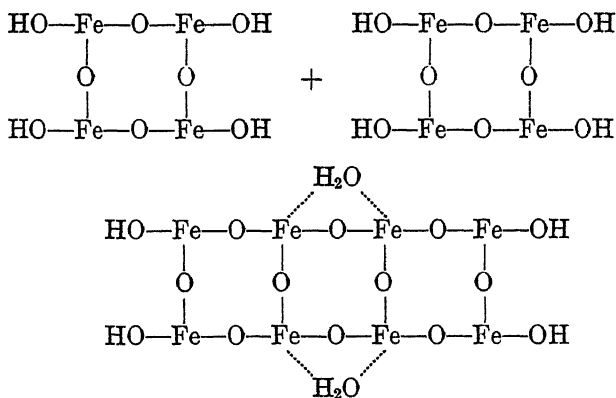
Here, again, all but one hydrogen may be replaced by silver, but the ferrite would have the composition

$$\text{Ag}_2\text{O}:\text{Fe}_2\text{O}_3 = 1:1.14$$

Evidently, as the degree of polymerization increases, the  $\text{Ag}_2\text{O}:\text{Fe}_2\text{O}_3$  ratios of the silver ferrites that can be formed will approach asymptotically to unity. It was estimated that in solutions of normal alkalinity, the degree of polymerization might be as high as 40 or 50. Like the ortho-hydroxide, poly-ortho-hydroxide is soluble in concentrated acetic acid and in 30 per cent nitric acid. This conclusion regarding polymerization was drawn from experiments on rapid aging in boiling alkaline media. The same type of polymerization occurs on slow aging at room temperature (57).

Beyond poly-ortho-hydroxide, aging may follow either of two paths. We shall consider, first, the path to goethite which is followed when the hydroxide concentration is between 0.01 *N* and 3.5 *N*. Krause and his coworkers have identified two intermediate phases along this path, amorphous ferric acid and polyferric acid (polyeisenige Säure) (59). This path is characterized by decreasing affinity for silver and by the fact that the members are insoluble in concentrated acetic acid and 30 per cent nitric acid. This last property requires a ring instead of a chain structure, because this insolubility is also a property of meta-hydroxide. Hence, Krause and Ciokówna (54) consider that ring closure occurs in poly-ortho-hydroxide to form large ring molecules. All of the hydrogen atoms are replaceable by silver. The product of this ring closure is called amorphous ferric acid. Its chemical reactions are identical with those of meta-hydroxide,  $\gamma$ -ferric oxide hydrate, but they differ in two respects. Amorphous ferric acid ages in the direction of goethite,  $\alpha$ -Fe<sub>2</sub>O<sub>3</sub>·H<sub>2</sub>O, while  $\gamma$ -ferric oxide hydrate is stable. And it does not yield ferromagnetic  $\gamma$ -ferric oxide on dehydration. This circumstance requires an explanation and a defense of the classification given by the silver analysis. It has been given by Krause and Lewandowski (57).

According to Krause, Lakósciukówna, and Cichowski (56) the next step on the way to goethite is aggregation of the large ring molecules into a network of polyferric acid. Schematically this aggregation may be represented by:



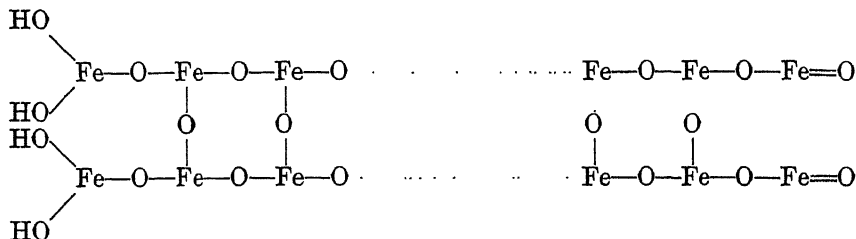
The amorphous ferric acid molecules are joined to each other by oxygen bridges, and the water becomes coördinatively bound. Obviously, as the network becomes large, all of the hydrogen appears in the water, so that polyferric acid, like goethite itself, does not form silver ferrite.

For a discussion of the complex mixtures that appear during the final



stage, crystallization to goethite, and of the difficulties in the way of separating them, the reader is referred to the original paper by Krause and Torno (59).

Under water or in a weakly alkaline medium,  $N < 0.01$ , aging from poly-ortho-hydroxide leads to hydrohematite. Along this path ring closure does not occur. The long chains of poly-ortho-hydroxide polymerize into a bundle of chains. Water is eliminated and the structure may be represented by:



It is called (59) poly-ortho-hydroxide-associate.

The final step along this path is crystallization of the aggregate of chain molecules to  $\alpha$ -ferric oxide and the zeolitic binding of the water. Krause has not as yet considered further aging from hydrohematite. It will be recalled that, according to Hüttig, hydrohematite eventually ages to goethite by way of limonite, as indicated in the diagram of figure 1.

Krause, Lakósciukówna, and Cichowski (56) found that both paths, ring structure to goethite and chain structure to hydrohematite, are followed when aging beyond poly-ortho-hydroxide is carried out at higher temperatures up to 150°C.

#### REFERENCES

- (1) ABEGG, R.: Z. Elektrochem. **10**, 77 (1904).
- (2) ABEGG: Handbuch der anorganischen Chemie, Eisen, 3rd edition, Vol. 4, Part 2B. S. Hirzel, Leipzig (1930).
- (3) ABRAHAM, H., AND PLANIOL, R.: Compt. rend. **180**, 1328 (1925).
- (4) ALBRECHT, W. H.: Ber. **62**, 1475 (1929).
- (5) ALBRECHT, W. H.: Private communication.
- (6) ALBRECHT, W. H., AND WEDEKIND, E.: Z. anorg. allgem. Chem. **202**, 205, 209 (1931).
- (7) BAUDISCH, O.: Ber. **54**, 406 (1921).
- (8) BAUDISCH, O.: Naturwissenschaften **16**, 542 (1928).
- (9) BAUDISCH, O.: Biochem. Z. **232**, 35 (1931).
- (10) BAUDISCH, O.: Science **77**, 317 (1933).
- (11) BAUDISCH, O.: Biochem. Z. **258**, 69 (1933).
- (12) BAUDISCH, O., AND ALBRECHT, W. H.: J. Am. Chem. Soc. **54**, 943 (1932).
- (13) BAUDISCH, O., AND MAYER, P.: Biochem. Z. **107**, 1 (1920).
- (14) BAUDISCH, O., AND WELO, L. A.: J. Biol. Chem. **61**, 261 (1924).

- (15) BAUDISCH, O., AND WELO, L. A.: J. Biol. Chem. **64**, 753 (1925).
- (16) BAUDISCH, O., AND WELO, L. A.: Naturwissenschaften **13**, 749 (1925); J. Biol. Chem. **65**, 215 (1925).
- (17) BAUDISCH, O., AND WELO, L. A.: Naturwissenschaften **21**, 659 (1933).
- (18) BAUDISCH, O., AND WELO, L. A.: Naturwissenschaften **21**, 593 (1933).
- (19) BÖHM, J.: Z. Krist. **68**, 567 (1928).
- (20) BRAGG, W. H.: Phil. Mag. [6] **30**, 305 (1915).
- (21) CHANDRA, H.: Inaugural Dissertation, Berlin, 1913.
- (22) CHAUDRON, G., AND FORESTIER, H.: Ann. chim. [9] **16**, 221 (1921); [10] **9**, 339 (1928); Compt. rend. **172**, 152 (1921); **178**, 2173 (1924).
- (23) CLAASSEN, A.: Proc. Phys. Soc. London **38**, 482 (1926).
- (24) DANA, J. D.: Am. J. Sci. Arts **44**, 89 (1867).
- (25) DAVEY, W. P.: Phys. Rev. [2] **21**, 716 (1923).
- (26) DEISS, E., AND SCHICKORR, G.: Z. anorg. allgem. Chem. **172**, 32 (1928).
- (27) DOELTER, C.: Kolloid-Z. **7**, 29, 86 (1910).
- (28) EMMETT, P. H., AND LOVE, K. S.: J. Phys. Chem. **34**, 41 (1930).
- (29) FOOTE, H. W., AND SAXTON, B.: J. Am. Chem. Soc. **38**, 588 (1916); **39**, 1103 (1917).
- (30) FORESTIER, H., AND CHAUDRON, G.: Compt. rend. **181**, 509 (1925).
- (31) GEILMANN, W., KLEMM, W., AND MEISEL, K.: Naturwissenschaften **20**, 639 (1932).
- (32) GMELIN: Handbuch der anorganischen Chemie, Eisen. Part 59B. Verlag Chemie, Berlin.
- (33) GOLDBERG: Dissertation, Dresden, 1914.
- (34) GOLDSCHMIDT, V. M.: Geochem. Verteil. der Elemente, Part VII, Oslo, 1926; Trans. Faraday Soc. **25**, 252 (1929).
- (35) GOLDSZTAUB, S.: Compt. rend. **193**, 533 (1931).
- (36) HABER, F.: Ber. **55**, 1717 (1922); Naturwissenschaften **13**, 1007 (1925).
- (37) HAHN, O.: Ann. **440**, 121 (1924).
- (38) HAHN, F., AND HERTRICH, F.: Ber. **56**, 1729 (1923).
- (39) HAUSKNECHT, P.: Inaugural Dissertation, Strassburg, 1913.
- (40) HEDVALL, J. A., AND SJÖMAN, P.: Z. Elektrochem. **37**, 30 (1931).
- (41) HEISENBERG, W.: Z. Physik **49**, 619 (1928).
- (42) HILPERT, S.: Ber. **42**, 2248 (1909); Ber. deut. physik. Ges. **11**, 293 (1909); J. Iron Steel Inst. London **82**, 65 (1910).
- (43) HUGGETT, J.: Ann. chim. [10] **11**, 447 (1929).
- (44) HÜTTIG, G. F.: Z. anorg. allgem. Chem. **114**, 161 (1920).
- (45) HÜTTIG, G. F.: Fortschr. Chem. Physik physik. Chem. **18**, 1 (1924).
- (46) HÜTTIG, G. F.: Z. Elektrochem. **37**, 631 (1931).
- (47) HÜTTIG, G. F., AND GARSIDE, H.: Z. anorg. allgem. Chem. **179**, 49 (1929).
- (48) HÜTTIG, G. F., AND KITTEL, H.: Z. anorg. allgem. Chem. **199**, 129 (1931).
- (49) HÜTTIG, G. F., AND MÖLDNER, H.: Z. anorg. allgem. Chem. **196**, 177 (1931).
- (50) HÜTTIG, G. F., RADLER, H., AND KITTEL, H.: Z. Elektrochem. **38**, 492 (1932); Naturwissenschaften **20**, 639 (1932).
- (51) HÜTTIG, G. F., AND ZÖRNER, A.: Z. anorg. allgem. Chem. **184**, 180 (1929).
- (52) HÜTTIG, G. F., AND ZÖRNER, A.: Z. Elektrochem. **36**, 259 (1930).
- (53) KRAUSE, A.: Z. anorg. allgem. Chem. **174**, 145 (1929); **195**, 228 (1931).
- (54) KRAUSE, A., AND CIOKÓWNA, M.: Z. anorg. allgem. Chem. **204**, 20 (1932).
- (55) KRAUSE, A., CZAPSKA, Z., AND STOCK, J.: Z. anorg. allgem. Chem. **204**, 385 (1932).

- (56) KRAUSE, A., LAKÓSCIUKÓWNA, H., AND CICHOWSKI, J.: *Z. anorg. allgem. Chem.* **208**, 282 (1932).
- (57) KRAUSE, A., AND LEWANDOWSKI, A.: *Z. anorg. allgem. Chem.* **206**, 328 (1932).
- (58) KRAUSE, A., AND PILAWSKI, K.: *Z. anorg. allgem. Chem.* **197**, 301 (1931).
- (59) KRAUSE, A., AND TORNÓ, H.: *Z. anorg. allgem. Chem.* **211**, 98 (1933).
- (60) KURNAKOW, N. S., AND RODE, E. J.: *Z. anorg. allgem. Chem.* **169**, 57 (1928).
- (61) LEFORT, J.: *Compt. rend.* **34**, 488 (1852).
- (62) LIVERSIDGE, A.: *Repts. Australian Assocn. Adv. Sci.*, p. 302 (1892).
- (63) LORIE, S.: *Inaugural Dissertation*, Zürich, 1912.
- (64) LUYKEN, W., AND BIERBRAUER, E.: *Mitt. Kaiser-Wilhelm Inst. Eisenforsch. Düsseldorf* **11**, 95 (1929).
- (65) MALAGUTI, F.: *Ann. chim. phys.* [3] **69**, 214 (1863).
- (66) MAURAIN, C.: *J. phys.* [3] **10**, 123 (1901).
- (67) MERCK, F., AND WEDEKIND, E.: *Z. anorg. allgem. Chem.* **186**, 49 (1930).
- (68) MITTASCH, A.: *Z. Elektrochem.* **41**, 827 (1928).
- (69) MIYAMOTO, S.: *Bull. Chem. Soc. Japan* **2**, 40 (1926).
- (70) MUCK, F.: *Jahresber.* **1867**, p. 287; *Z. Chem.* [2] **4**, 41 (1868).
- (71) NATTA, G., AND CASSAZA, E.: *Atti accad. Lincei* [6] **5**, 807 (1927).
- (72) NISHIKAWA, S.: *Proc. Math. Phys. Soc. Tokyo* **8**, 199 (1915).
- (73) PAULING, L.: *J. Am. Chem. Soc.* **51**, 1010 (1929).
- (74) PLÜCKER, J.: *Pogg. Ann.* **74**, 321 (1848).
- (75) POSNJAK, E., AND MERWIN, H. E.: *Am. J. Sci.* [4] **47**, 311 (1919).
- (76) QUARTAROLI, A.: *Gazz. chim. ital.* **46**, 219 (1916).
- (77) ROBBINS, J.: *Chem. News* **1**, 11 (1859).
- (78) RODT, V., AND MECKLENBURG, W.: *Z. anorg. allgem. Chem.* **182**, 130 (1928).
- (79) RUFF, O.: *Ber.* **34**, 3417 (1901).
- (80) SACHSE, H.: *Z. physik. Chem.* **9B**, 83 (1930).
- (81) SACHSE, H., AND HAASE, R.: *Z. physik. Chem.* **148A**, 401 (1930).
- (82) SCHIFF, H.: *Ann.* **114**, 199 (1860).
- (83) SCHWARZ, F. K. T.: *Klin. Wochschr.* **8**, 923 (1929).
- (84) SERRES, A.: *Ann. phys.* [2] **17**, 5 (1932).
- (85) SIMON, A.: *Über die Konstitution und den stabilen Endzustand von Hydrogelen. Festschrift der Technischen Hochschule, Stuttgart*, 1929.
- (86) SIMON, A., AND SCHMIDT, T.: *Kolloid-Z., Zsigmondy Festschrift*, **36**, 65 (1925).
- (87) SIMON, A., AND THALER, E.: *Z. anorg. allgem. Chem.* **161**, 113 (1927).
- (88) SMEKAL, A.: *Z. Elektrochem.* **35**, 567 (1929); *Z. angew. Chem.* **42**, 489 (1929).
- (89) SOSMAN, R. B.: *J. Wash. Acad. Sci.* **7**, 55 (1917).
- (90) SOSMAN, R. B., AND POSNJAK, E.: *J. Wash. Acad. Sci.* **15**, 329 (1925).
- (91) STOLLENWERK, W., AND VON WRANGELL, M.: *Z. Elektrochem.* **33**, 501 (1927).
- (92) TAYLOR, W. H.: *Z. Krist.* **72**, 205 (1929).
- (93) THEWLIS, J.: *Phil. Mag.* [7] **12**, 1089 (1931).
- (94) TOMMASI, D.: *Ricerche sulle formole di constitutione dei composti ferrici. Parte I; Idrati ferrici*, Firenze, 1879; *Ber.* **12**, 1929, 2334 (1879); *Moni. sci.* [4] **2**, 164 (1879); *Bull. soc. chim.* [2] **38**, 152 (1882).
- (95) TOMMASI, D.: *Moni. sci.* **31**, 164 (1888).
- (96) TOMMASI, D., AND PELLIZZARI, G.: *Bull. soc. chim.* [2] **37**, 196 (1882).
- (97) TWENHOFEL, L. H.: *Econ. Geol.* **22**, 180 (1927).
- (98) VAN BEMMELEN, J. M.: *Rec. trav. chim.* **7**, 37 (1888); *Z. anorg. allgem. Chem.* **20**, 185 (1892); **22**, 313 (1900).

- (99) VAN BEMMELN, J. M., AND KLOBBIE, E. A.: J. prakt Chem. **46**, 497 (1892).
- (100) VON KOBELL, F.: Grundzüge der Mineralogie, p. 304. Nürnberg (1838).
- (101) WEDEKIND, E., AND FISCHER, H.: Ber. **60**, 541 (1927).
- (102) WELO, L. A., AND BAUDISCH, O.: Phil. Mag. [6] **50**, 399 (1925).
- (103) WELO, L. A., AND BAUDISCH, O.: Phys. Rev. [2] **25**, 587 (1925); Abstract of paper read before the American Physical Society, February 27-28, 1925.
- (104) WELO, L. A., AND BAUDISCH, O.: Phil. Mag. [7] **3**, 396 (1927); Naturwissenschaften **14**, 1005 (1926).
- (105) WELO, L. A., AND BAUDISCH, O.: Am. J. Sci. In press.
- (106) WELO, L. A., AND BAUDISCH, O.: Phil. Mag. [7] **17**, 753 (1934).
- (107) WILLIAMS, R. D., AND THEWLIS, J.: Trans. Faraday Soc. **27**, 767 (1931).
- (108) WILLSTÄTTER, R., AND KRAUT, H.: Ber. **56**, 149, 1117 (1923); **57**, 1082 (1924); **59**, 2541 (1926).
- (109) WILLSTÄTTER, R., KRAUT, H., AND ERBACHER, O.: Ber. **58**, 2448, 2458 (1925).
- (110) WILLSTÄTTER, R., KRAUT, H., AND FREMERY, W.: Ber. **57**, 1491 (1924).
- (111) WILLSTÄTTER, R., KRAUT, H., AND LOBINGER, K.: Ber. **58**, 2462 (1925).
- (112) ZSIGMONDY, R.: Lehrbuch der Kolloidchemie, 5th edition. Leipzig (1925).



# THE CHEMICAL NATURE OF ENZYMES

HENRY TAUBER

*Department of Physiology and Physiological Chemistry, New York Homeopathic Medical College and Flower Hospital, New York City*

*Received May 29, 1934*

Efforts to elucidate the chemical nature of enzymes date back many years, and the results of these earlier investigations, as well as those of today, may be grouped into two classes. One indicates that enzymes are of protein nature, and the other attempts to show that these same enzymes are non-protein substances. It will be seen that many of the earlier as well as recent reports have been the basis of erroneous interpretations, and that it is quite clear now that many enzymes are proteins.

## THE CHEMICAL NATURE OF PEPSIN

Pepsin is the proteolytic enzyme secreted by certain cells of the gastric mucosa of adult mammals. It is characteristic in being active only in a distinctly acid medium. In the mucosa, however, it is present as a precursor, an inactive form, called pepsinogen or propepsin. Pepsin digests proteins to the proteose-peptone stage. Like other proteases, it splits peptide linkages.

In 1885 Sundberg (76) prepared a pepsin solution which gave negative protein color tests. He extracted the mucosa of the calf's stomach with saturated sodium chloride solution for a few days and autolyzed the extract for about two weeks. Then he precipitated the pepsin with calcium phosphate and redissolved it in hydrochloric acid. After dialysis, this pepsin solution lost most of its activity, but gave no protein color tests. According to Sundberg, however, his pepsin gave a very strong odor, when heated on platinum, resembling that of burning horn. Similar results were reported in 1861 by Brucke (6).

In 1902 Pekelharing (61) published a very interesting paper on pepsin. He purified gastric juice and acid extracts of the gastric mucosa of the pig by dialysis and found that the pepsin separated in refractive globules. These preparations of Pekelharing were very active.

Northrop (54) found that the precipitate which formed in the dialyzing sac when the procedure of Pekelharing was followed, appeared in more or less granular form and filtered rather easily, as though it were on the verge

of crystallization, and that "this precipitate dissolved on warming the suspension and it was eventually found that it could be induced to crystallize by warming to 45°C., filtering, and allowing the filtrate to cool slowly." Northrop has succeeded in working out a method by which pepsin can be obtained in crystalline form in large quantities. Precipitation by magnesium sulfate in acid medium and solution in alkali with subsequent acid precipitation is the principle of his method. Northrop has recrystallized pepsin seven times and has found by chemical and physical tests that the enzyme is a protein. A typical experiment is the heat inactivation of

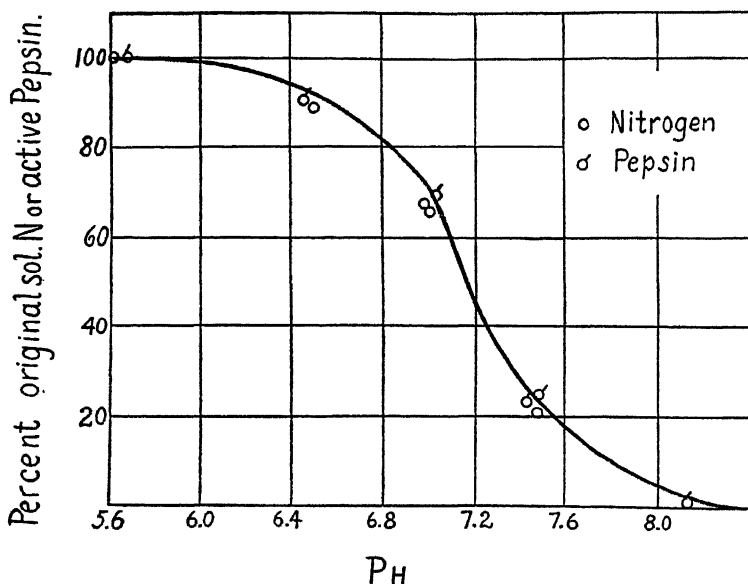


FIG. 1. PERCENTAGE INACTIVATION AND PERCENTAGE DENATURATION OF CRYSTALLINE PEPSIN AT VARIOUS pH VALUES AT 20°C.

(Northrop: J. Gen. Physiol. 13, 739 (1930))

pepsin. The latter parallels the denaturation of the protein. Percentage activity left after heat inactivation is the same as percentage of soluble nitrogen remaining (see figure 1). Northrop (57) has also obtained crystalline pepsin from the bovine gastric juice.

According to Levene and Helberger (41) crystalline pepsin is in the main an individual protein, but it probably contains a small admixture of an extraneous protein.

A separation of crystalline pepsin into two fractions having different gelatin-liquefying power has been reported by Holter (20). That crystalline pepsin contains a small amount of an enzyme with high gelatin-liquefying power was first stated by Northrop (55).

Loughlin (45) has prepared crystalline pepsin by Northrop's method. This product was identical with that of Northrop, both in general properties and in specific proteolytic activity.

Recently, Willstätter and Rohdewald (98) verified the findings of Sundberg and those of Brucke on the preparation of a protein-free pepsin solution. They, too, state that the activity of their solution decreased considerably in the course of preparation, resulting in a very slightly active *solution*, but that this solution gave no protein color tests. Tauber (78) has shown that these findings of Willstätter and Rohdewald were based on an erroneous interpretation, since active enzyme solutions may be too dilute to give positive protein color tests. As to Northrop's crystalline pepsin, Willstätter and Rohdewald (98) do not believe that these crystals are identical with the enzyme pepsin. Willstätter has been a staunch supporter of the carrier theory. Generally by a "carrier" is meant any colloid of high molecular weight, such as a protein or a carbohydrate which carries the active enzyme. This carrier is said to be exchangeable for other carriers. The enzyme itself, according to this theory, is a substance which cannot be classified with any of the known chemical compounds. However, Willstätter and Rohdewald have now modified the carrier theory by saying that enzymes have a "necessary colloidal carrier," and that the chemical nature of the *latter* is still unknown.

Glutathione acts in every way like an enzyme, and if it were inactivated upon heating its aqueous solution, it would be classified as an enzyme. In that case, one might say that the tripeptide is only the carrier of the active principle. Colloidal platinum has an enormous catalytic effect as compared to solid platinum. Is the colloidal platinum a carrier of an active substance or principle? It should be noted that the carrier theory is not new. J. Perrin (62) proposed it as early as 1905, and Mathews and Glenn (48) wrote in favor of it in 1911.

Dyckerhoff and Tewes (9) and Waldschmidt-Leitz and Kofranyi (90) published experiments purporting to show that crystalline enzymes are really crystalline proteins carrying the non-protein enzymes. They have compared Northrop's crystalline pepsin with edestin and other crystalline plant proteins which have been in suspension in pepsin solution. These proteins were found to be just as active as the crystalline pepsin, and accordingly these workers maintain that the active principle has exchanged its carrier. Northrop (58) repeated the work of these authors and found that, "The pepsin protein is taken up as such and the quantity of protein taken up by the foreign protein is just equivalent to the peptic activity found in the complex." Results similar to those of Northrop were reported by Sumner (71), who absorbed crystalline pepsin on casein and ovalbumin.

Dyckerhoff and Tewes stated that rennin acts in a similar manner, i.e..



it may be absorbed on crystalline edestin in active form, exchanging its carrier. However, they do not furnish experimental evidence concerning the absorption of rennin. Tauber and Kleiner (84), interested especially in the nature of rennin, have attempted to prepare "crystalline edestin-rennin," using methods similar to those of Dyckerhoff and Tewes and of Waldschmidt-Leitz and Kofranyi. Preparations of only slight activity were obtained, indicating that no exchange of the protein carrier took place.

#### THE RENNIN-PEPSIN CONTROVERSY AND THE CHEMICAL NATURE OF RENNIN

The controversy as to whether rennin and pepsin are separate enzymes is of early date. Hammarsten on the one hand represented those who believed in the existence of a separate enzyme, rennin having only milk-clotting power; Pavlov, on the other hand, was the leader of a school which maintained the existence of a mammoth molecule with side chains acting proteolytically at an acid pH and rennetically at the neutral point. The latter view has been supported by the fact that all proteases are able to clot milk.

According to Linderström-Lang (1928) the casein of the milk is a system of three components, one of which acts as a protective colloid for the other two. Any proteolytic enzyme has the ability to attack the protective colloid and under proper conditions (presence of calcium, pH, etc.) coagulation of the casein takes place. The clot is called calcium paracaseinate, and is insoluble. An attack on the casein, however, takes place also in the absence of calcium. Then paracasein, a soluble complex, forms, which subsequently may be precipitated on the addition of soluble calcium salts. In contrast to the very acid pH required for peptic activity (i.e., the digestion of other proteins by pepsin), pepsin clots milk at a practically neutral pH. It is not known whether the milk-clotting activity of rennin is a proteolytic process or not, i.e., whether opening of peptide linkages takes place.

Luers and Bader (46) have recently reported an attempt to separate rennin from pepsin using the adsorption method. These authors found their preparation to be nearly free of protein, containing only 0.678 per cent nitrogen, possessing high rennet activity as well as relatively high peptic activity. They state that their results do not allow them to decide in the favor of two distinct enzymes, rennin and pepsin.

Edie (10) obtained a pepsin preparation free of milk-coagulating power.

Tauber and Kleiner (81) have been able to isolate by isoelectric precipitation an extremely active rennin preparation, from fresh fourth stomachs of the calf. This possessed the highest activity so far reported. The minced mucosa was extracted for 8 minutes with 0.04 *N* hydrochloric acid.

The clear filtrate was dialyzed against distilled water until the pH was about 5.4. Then alcohol was added up to 50 per cent, which precipitated the rennin. The precipitate was dissolved in water freed from insoluble matter and precipitated as before. This procedure was repeated several times.

For peptic activity determination, the formol titration method was employed, and the increase of nitrogen not precipitated by 10 per cent trichloroacetic acid was measured by the micro-Kjeldahl method. Coagulated egg albumin of pH 2.0 was used as a substrate. This rennin contained no pepsin, since it did not produce an increase in formol titration and only a negligible increase in nitrogen not precipitated by trichloroacetic acid. It clotted 4,550,000 times its weight of milk (pH 6.2) in 10 minutes at 40°C. This is more than 100 times as active as Parke, Davis and Company's 1:30,000 preparation, and 2000 times as active as the original extract of the calf's mucosa. Crystalline pepsin has less milk-clotting power than concentrated rennin.

The vacuum-dried rennin differs chemically from crystalline pepsin in many ways. It is very readily soluble in slightly acidified water; it is not coagulated by heat, whereas pepsin is. It gives a pink biuret test, whereas pepsin gives a violet color. The Millon and Hopkins-Cole tests are not given. In contrast to pepsin, rennin diffuses quickly through dialyzing membranes. The isoelectric point of rennin is at 5.4, whereas that of pepsin is at 2.7. Pepsin contains phosphorus, whereas rennin does not. *Pepsin is an albumin; rennin a thioprotease.* These results of Tauber and Kleiner on the purification of rennin have been verified by Holter (21) working in the laboratory of Sørensen.

Thus, it is now obvious that rennin, the protease of the young mammal's gastric mucosa, and pepsin, the protease of the adult mammal's gastric mucosa, are distinct entities. It should be noted, however, that the precursors of these enzymes and not the active enzymes are present in the mucosa. (See also below.) Rennin is a milk-clotting enzyme; pepsin a typical protein-hydrolyzing enzyme with considerable milk-coagulating power.

#### THE CHEMICAL NATURE OF TRYPSIN

Trypsin as obtained from and secreted by the pancreas is almost completely inactive, and as such it is called trypsinogen (precursor). It can be activated, however, to an extremely powerful enzyme, if an activator, called "enterokinase" and prepared from the mucosa of the small intestine, is added to it. Trypsinogen, however, becomes active ("autoactivation") without the addition of enterokinase on standing at room temperature for several days. This may be due to activation by slightly active tissue

proteases. Trypsin digests practically all proteins, including proteoses and peptones, to polypeptides. The polypeptides are digested by specific peptidases, produced mainly by the intestinal mucosa. The optimum pH of trypsin, as also of all other enzymes, varies slightly with the substrate used. It is at about 8 to 9. Trypsin is also active in slightly acid solution.

A crystalline protein possessing tryptic properties has been obtained by Northrop and Kunitz (56, 59) from beef pancreas. A large number of experiments have been carried out, which showed that the protein crystals are identical with the proteolytic enzyme. This trypsin does not become more active on the addition of enterokinase. It is completely inactive towards peptides, has a slight milk-clotting power, an isoelectric point of about 7.0, and an optimum pH of 8.0 to 9.0 when casein is used as a sub-

TABLE 1

*Showing that trypsin can clot milk only within a certain range of concentration*

Tauber and Kleiner (85)

TRYPSIN	RESULTS AFTER 10 MINUTES	TRYPSIN	RESULTS AFTER 10 MINUTES
<i>per cent</i>		<i>per cent</i>	
2.00	No clot	0.10	Slight clot
1.00	No clot	0.08	Heavy curd
0.75	No clot	0.05	Heavy curd
0.50	Slight clot	0.03	No clot
0.20	Slight clot		

To 10-cc. samples of calcium chloride-milk, 1 cc. of each of the various trypsin solutions was added. The pH of this milk was 6.4, as determined by the electro-metric method. Temperature, 40°C.

After 5 hours results remained the same.

strate. The milk-clotting of this trypsin, according to Northrop and Kunitz, is due to a contamination of the trypsin by a milk-coagulating enzyme. Tauber and Kleiner (85), however, have shown that the trypsin itself can clot milk, but only within a certain limited range of concentration; if too dilute or if too concentrated, no coagulation will occur (see table 1). When concentrated trypsin solutions, like those used in protein digestion experiments, are employed, trypsin changes the casein molecule so rapidly beyond the paracasein stage, that the milk will not clot even after the subsequent addition of a very active rennin solution (see table 2). Thus, Tauber and Kleiner have definitely established that the milk-clotting power is a function of the trypsin molecule. They also found that the velocity of milk coagulation in all three cases—rennin, pepsin, and trypsin—is proportional to the hydrogen-ion concentration, and if the milk is adjusted to a low pH so that the proteolytic activity of the (concentrated) trypsin is depressed, coagulation may occur (see table 3).

Kleiner and Tauber have recently obtained from pancreatic tissue, by complete autolysis (eighteen months), dialysis, and acetone precipitation,

TABLE 2  
*Showing rapid effect of trypsin on casein*  
Tauber and Kleiner (85)

EXPERIMENT NO.	PERCENT OF TRYPSIN SOLUTION	RESULTS AFTER 10 MINUTES	AFTER 10 MINUTES OF TRYPTIC DIGESTION 1 CC. RENNIN SOLUTION CONTAINING 13,000 UNITS* PER CC. WAS ADDED	REMARKS
1	0.015	No clot	Clot in 8 min.	Trypsin too dilute to affect casein molecule (or complex)
2	0.030	No clot	Clot in 7 min.	
3	0.300	Clot	No rennin added	At this pH (6.4) and concentration trypsin readily clots milk
4	0.600	No clot	No clot in 5 hrs.	Trypsin solutions of these concentrations change casein to such degree that insoluble Ca paracaseinate cannot form, even on addition of active rennin
5	1.000	No clot	No clot in 5 hrs.	
6	4.000	No clot	No clot in 5 hrs.	

To 10 cc. of calcium chloride-milk of pH 6.4, 1 cc. of the trypsin solution was added. Temperature, 40°C.

\* A unit is the amount of milk (pH 6.2) in mg. clotted by 1 mg. of rennin.

TABLE 3  
*Milk-clotting activity of rennin, pepsin, and trypsin at various pH values*  
Tauber and Kleiner (85)

pH	CLOTTING TIME OF RENNIN	CLOTTING TIME OF PEPSIN	CLOTTING TIME OF TRYPSIN
	<i>minutes</i>	<i>minutes</i>	<i>minutes</i>
5.6	3.5	1.5	3.0
5.9	4.5	2.0	4.0
6.1	8.0	3.0	6.0
6.4	10.0	11.0	No clot in 40 min.*
6.6	29.0 very slight	No clot in 40 min.	No clot in 40 min.*
6.8	36.0 very slight	No clot in 40 min.	No clot in 40 min.*

The pH of the milk was adjusted with hydrochloric acid or sodium hydroxide and electrometrically controlled. To 10-cc. samples of milk, 0.5 cc. of each of the enzyme solutions was added. The trypsin solution was 1 per cent. The pepsin solution contained 18,000 rennet units per cubic centimeter. The rennin solution had an activity of 16,000 units per cubic centimeter. Temperature, 40°C.

\* No clot occurred after adjustment to pH 5.6.

a trypsin preparation which does not resemble a protein of the ordinary type. The activity of this preparation is about the same as that of crystal-

line trypsin. The dry enzyme is very soluble in water and dilute alkali, slightly soluble in dilute mineral acids, and insoluble in organic solvents. It dialyzes readily.

The protein color tests as obtained with a 1 per cent solution were as follows: Xanthoproteic test, slightly positive; Folin-Denis test, positive; various modifications of biuret test, negative; Millon's test, negative; heat coagulation test, positive only on prolonged boiling, when a slight precipitate was obtained. The isoelectric point is at pH 6.2. This trypsin may be either a broken-down product of protein, or a protease differing from that obtained directly from fresh pancreatic tissue. The elementary composition differs considerably from that of crystalline trypsin. The milk-clotting power is about one-thousandth that of rennin. The milk-coagulating power is a function of the trypsin molecule.

#### CHYMOTRYPSIN

An interesting crystalline proteolytic enzyme has been recently isolated by Kunitz and Northrop (38) from fresh beef pancreas. The enzyme as first obtained is in an inactive form and cannot be activated by enterokinase. This inert form is called chymotrypsinogen. If, however, a very small amount of active trypsin is added, the inactive chymotrypsin becomes very active. The gelatin-hydrolyzing power of this enzyme is less than that of trypsin, but its milk-clotting power is greater. The chymotrypsinogen has been recrystallized ten times with unchanged optical activity and constant proteolytic activity when activated. The preparations are pure proteins and "the proteolytic activity is a property of the protein molecule."

Thus another enzyme has been obtained in crystalline form and another controversy has been settled, since it had previously been uncertain whether trypsinogen may be activated by a small amount of active trypsin.

#### THE CHEMICAL NATURE OF PAPAIN AND CATHEPSIN

Papain is a plant protease obtainable from the fruit or milky juice of the melon tree (*Carica papaya*). Its optimum pH is about 7.0. Cathepsin is an animal tissue protease, having also an optimum pH of about 7.0.

Papain is not completely active when prepared from the plant material, but is readily activated with hydrogen sulfide. Bersin and Logenmann (4) explain the activation of papain by reduced glutathione as a reduction of the oxidized form of the enzyme having the S—S linkages. They inactivated papain with hydrogen peroxide and reactivated it with hydrogen sulfide or with other reducing agents such as sodium sulfite and reduced glutathione. A possibility of an active hydrogen group in the enzyme is suggested, since diazomethane inactivates the enzyme.

Bersin (3) offered further evidence that the activation of a glutathione-free papain is due to the reduction of the S—S groupings present in the enzyme-protein complex to SH. He activated papain by a number of reducing agents such as hydrogen sulfide, glutathione, hydrogen cyanide, sulfur dioxide, cysteine, and succinate plus dehydrogenase. Ultra-violet irradiation likewise produced an activation which may be explained by the recent observation that cysteine was formed from cystine by similar treatment. Organic arsenic compounds and such inorganic arsenic compounds as have the ability to reduce the S—S group to the SH group, also activate papain. Ascorbic acid, however, could not activate papain, which Bersin explains to be due to the fact that it does not form thiol derivatives from disulfides. Contrary to Waldschmidt-Leitz (89), Bersin could not activate phosphatase with glutathione. Purr (64) activated cathepsin, using a carcinoma-glycerol suspension, by ascorbic acid with practically the same degree of activation as was obtained when cysteine was used as an activator. He also activated liver arginase with ascorbic acid; ascorbic acid-iron was found to be the best activator.

Karrer and Zehender (28) freed liver cathepsin from its natural activators and were likewise able to reactivate it with ascorbic acid. The degree of activation was of the same magnitude as that of hydrogen sulfide activation.

These findings are interesting, since papain cannot be activated by ascorbic acid. From a physiological point of view, therefore, ascorbic acid appears to be more than only an antiscorbutic. Combined with glutathione, it catalyzes important oxidation and reduction systems, which may be reversibly oxidized and reduced establishing equilibrium between the S—S and SH compounds. Without these combined effects, certain intracellular enzymes would not be able to carry out their metabolic function, as indicated by the experiments of Purr and those of Karrer and Zehender.

#### THE CHEMICAL NATURE OF UREASE

Urease is found in large amounts in the seed of *Leguminosae*, and has been obtained in small amounts from various animal organs as well as from certain bacteria. The best sources of urease are jack bean and soy bean. It hydrolyzes urea into ammonia and carbon dioxide, and has an optimum pH of 7.

Urease was the first enzyme to be obtained in crystalline form. The rather simple method for the crystallization of this enzyme was described by Sumner in 1926 (70). The principle of his procedure is as follows: Jack bean meal is extracted for a short time with 30 per cent acetone and left to filter overnight in a refrigerator. Octahedric, microscopic crystals which

separate may be recrystallized from either dilute acetone or dilute alcohol. The activity of the crystals is considerable. The yields of urease crystals are small, however. Since the method of crystallization is very convenient, it should be used whenever a pure preparation of urease is desired. Sumner's results on the preparation and activity of crystalline urease were confirmed by Tauber in 1930 (77).

Crystalline urease gives all the protein color tests and the test for unoxidized sulfur. It is a globulin having an isoelectric point of about pH 5.0 (72).

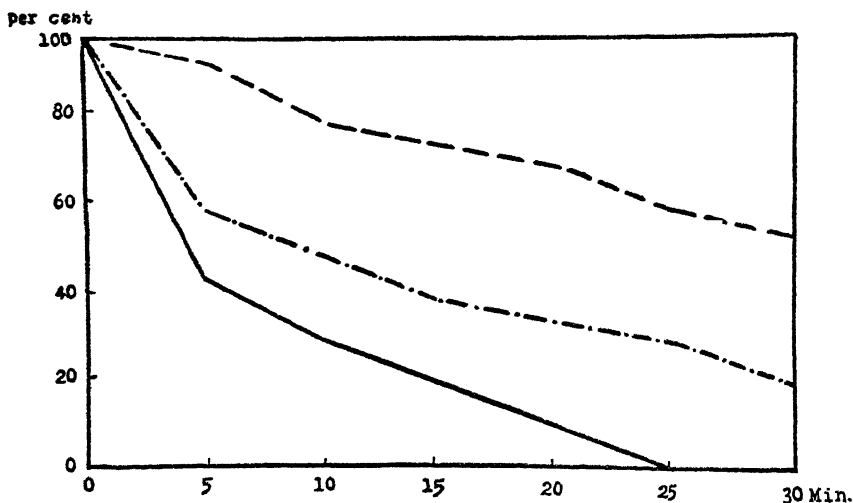


FIG. 2. EXPOSURE OF CRYSTALLINE UREASE SOLUTIONS TO ULTRA-VIOLET RAYS

— represents 5 cc. of urease, without eosin, at 41°C. exposed at 25 cm.;  
 ----- represents 5 cc. of urease, without eosin, at 30°C. exposed at 50 cm.; - · - · -  
 represents the same plus 0.05 cc. of 0.01 per cent eosin. The function of the eosin is that of a photodynamic agent. (Tauber: J. Biol. Chem. 87, 625 (1930).)

A controversy as to the protein nature of urease between Waldschmidt-Leitz and Steigerwald (92) and Sumner and coworkers (73, 74) appears to be settled in favor of the latter workers, i.e., crystalline urease is a protein.

Ultra-violet irradiation inactivates crystalline urease (figure 2), which is apparently due to a molecular rearrangement of the protein molecule (77).

Hotchkiss and Tauber (22) have shown that crystalline urease solutions produce agglutination of erythrocytes. Agglutination did not inactivate the urease. With whole blood from dead rabbits no agglutination could be demonstrated, since serum tends to inhibit this reaction. Crystalline urease was found to be a highly purified form of a toxic hemagglutinin. The minimum effective dose was 0.00014 mg. of the enzyme in 1 cc. The

hemagglutinin obtained from the navy bean by Goddard and Mendel (15) was effective with a minimum dose of 0.0006 mg. in a total volume of 1 cc.

The toxicity of urease when injected into the animal body was ascribed to various causes. The results of Tauber and Kleiner (79), who also studied the toxicity of ammonium carbonate, point directly to ammonia poisoning, as was also suggested by Carnot and his coworkers (8) and by Lövgren (42). The minimum lethal dose for mice was determined by Kleiner and Tauber, and was found to be 0.09 Sumner units of urease per gram of body weight.

An immunity to urease was acquired by the experimental animals (79), and Sumner (30) described a method for the preparation of antiurease. They also found urease poisoning to be the result of ammonia intoxication. According to Howell (24) the hen is not poisoned by urease since it has only 2 mg. of urea per 100 cc. of blood; antiurease is produced, however.

Jacoby (27) studied the effect of cupric chloride, mercuric chloride, and silver nitrate on urease and its reactivation with potassium cyanide. At low temperatures the inactive urease-metal compound can be completely reactivated with potassium cyanide; if, however, the temperature is raised to 37°C., the inactivation is irreversible. Sumner, Lloyd, and Poland (75) recrystallized urease several times and found that it gave the same amount of red color in the nitroprusside test, indicating a sulfhydryl group in the urease molecule. According to these authors, this is the reason for the rapid destruction of urease by silver, mercury, and copper ions and by quinone and other oxidizing agents. The protection of urease by sulfhydryl compounds is thus explained.

Hellerman, Perkins, and Clark (19) subjected urease to oxidation and reduction. Urease solution, completely inactivated (by cuprous oxide), was reactivated by hydrogen sulfide to 83 per cent. Benzyl mercuric chloride inactivated a urease solution completely. An excess of hydrogen sulfide effected complete reactivation. If urease solutions are aerated in the presence of  $\text{Cu}^{++}$  rapid destruction of the enzyme takes place; sulfhydryl groups, however, cause reactivation. When urease is inactivated by iodine, potassium cyanide does not restore activity, but hydrogen sulfide does. The authors offer the following explanation for the above results: "... the more drastic action of iodine might involve a deep seated change due to a permanent substitution in the protein molecule as well as an oxidation of thiol sulfur beyond the dithio stage, in which case irreversible inactivation of a labile, highly specific protein enzyme molecule would be anticipated. On the other hand, more superficial substitutions in the protein molecule accompanying the oxidation of thiol to dithio, might be wholly reversed by such a reagent as  $\text{H}_2\text{S}$  but not necessarily by  $\text{KCN}$ ."

Similar studies of active groups in enzyme molecules and their inactiva-



tion and reactivation may throw light on the constitution of enzymes. Extensive investigations of such a nature are highly suggestive.

#### THE CHEMICAL NATURE OF THE YELLOW OXIDATION ENZYME

Warburg and Christian (93) prepared oxidation enzymes from various sources such as anaerobic bacteria, chlorella, and bottom yeast. Unlike the pheohemin of aerobic cells, oxidation enzymes are not poisoned by carbon monoxide or hydrogen cyanide. The oxidation enzyme of bottom yeast has been obtained in pure form and extensively studied by Warburg and Christian. It does not dialyze through cellophane and is destroyed when heated above 60°C. The aqueous solutions of the enzyme have a yellow-red color. The color disappears if treated with reducing agents, but it reappears if the solution is shaken with oxygen. The reduced form does not absorb in the visible spectrum and is called the "leuco form." The oxidation of the leuco enzyme is much quicker when methylene blue is used instead of molecular oxygen, and by titration with the former the concentration of the active group may be determined. The "enzyme" has the ability to transfer oxygen to Robison's hexosemonophosphate ester, the colored form being reduced by this process. Molecular oxygen reoxidizes the leuco compound *by direct chemical reaction*. *Since, like hemoglobin, this oxidation enzyme combines chemically with oxygen, it is obvious that it cannot be called an enzyme.* The reduction of the colored compound by hexosemonophosphate, however, is, according to Warburg and Christian, a catalytic effect. It requires a second enzyme and a coenzyme. The hexosemonophosphate, the second enzyme, and the coenzyme, are colorless and are called the "reducing system."

The coenzyme may be prepared from red blood cells and when purified until it is free of protein, it does not give a precipitate with trichloroacetic acid. It should be noted, however, that certain derived proteins, like peptones, do not give a precipitate with the latter. The coenzyme is sensitive to acid and alkali, but is stable in neutral solution.

The second enzyme (the enzyme of the reducing system) may be obtained from baker's yeast extracts prepared according to Lebedew's method. This enzyme does not dialyze through cellophane membranes. It is destroyed when heated at 60°C. for 15 minutes.

The oxidizing agents in the aerobic cell are molecular oxygen and ferric iron of pheohemin. The ferric iron reacts more rapidly with the leuco form of the colored pigment than the molecular oxygen, and consequently there is no direct reaction between molecular oxygen and the leuco form of the yellow-red dye, and the oxidation follows this course:  $O_2 \rightarrow$  pheohemin

(Fe<sup>++</sup>)

$\rightarrow$  pheohemin  $\rightarrow$  leuco form of dye  $\rightarrow$  pigment  $\rightarrow$  "reducing system" (I).  
(Fe<sup>+++</sup>)

If, however, the pheohemin is inactivated by carbon monoxide or hydrogen cyanide then the course is the following:  $O_2 \rightarrow$  leuco form  $\rightarrow$  pigment  $\rightarrow$  "reducing system" (II). Generally a decrease in the rate of oxidation takes place, if I is replaced by II.

Warburg and Christian (93) state that they have separated the yellow-red oxidation enzyme into a crystalline pigment (enzyme) "component" and a "protein carrier." Both, however, were obtained in an inactive form. They report: "In dem Ferment ist der Farbstoff an Eiweiss gebunden. Man kann Farbstoff-und Eiweiss Komponente trennen, indem man das Ferment bei 38° mit Methanol-Wasser (1 Vol. Wasser, 3 Vol. Methanol) schüttelt. Dann wird dass Eiweiss denaturiert, der Farbstoff geht in Lösung. Die eiweissfrei Lösung des Farbstoffs fluoresziert grün. Sie ist *katalytisch unwirksam*, aber in Bezug auf ihr Absorptionsspectrum (im sichtbaren) der Fermentlösung sehr ähnlich."<sup>1</sup>

#### THE CHEMICAL NATURE OF LIPASE

Pancreatic lipase is secreted by the pancreas. It hydrolyzes fats into fatty acids and glycerol. A specific lipase is present in the gastric mucosa, which hydrolyzes only emulsified fats such as the butter fat of milk and the fat of the egg yolk. Plants, too, contain lipases; *Ricinus communis* is one of the best sources.

Glick and King (14) used 10 per cent sodium chloride for the extraction of lipase from dried pancreas tissue and obtained more powerful preparations than were obtained by solvents used in earlier work. The lipase may be salted out quantitatively by saturation with magnesium sulfate. Magnesium salts activate the lipase. Expression of enzyme activity in units per milligram of nitrogen is suggested. The purified enzyme shows the general characteristics of a globulin.

Bamann and Laeverenz (2) described an accidental observation of a "crystalline lipase protein." They found that if a certain dry lipase preparation was extracted with 25 per cent glycerol and allowed to filter slowly for 21 days, active colorless needles, 2 mm. long, separated from the dark yellow filtrate. The chemical nature of this crystalline lipase preparation has not been further studied by these authors.

#### THE CHEMICAL NATURE OF PANCREATIC AMYLASE

Amylases are starch- and glycogen-splitting enzymes, and the end-product of hydrolysis is maltose. Saliva, the pancreas, and the liver are good sources of animal amylase. In plants, the germinating seed contains large amounts of this enzyme. The optimum pH of amylases differs

<sup>1</sup> For an extended discussion of the respiratory enzymes see H. von Euler's recent book,—Die Katalasen und die Enzyme der Oxydation und Reduction. J. Springer, Berlin (1934).

greatly; for instance, pancreatic amylase has an optimum pH of about 7, salivary amylase of 6.0, and malt amylase of about 5.0. These enzymes are peculiar in being inactive in the absence of neutral salts.

A crystalline protein with high amylolytic activity has been obtained from pancreatin by Caldwell, Booher, and Sherman (7). Waldschmidt-Leitz and Reichel (91), however, state that they have obtained an active protein-free amylase from pancreatic tissue. Salivary amylase is digested by trypsin, pointing to the protein nature of this amylase (Tauber and Kleiner).

Sherman, Caldwell, and Doebbeling (67) have purified malt amylase, which yields products of much higher activity than any previously reported. These authors attempted to purify their enzyme by means of adsorption by alumina gel and subsequent elution. This method was introduced at various stages in the purification process. However, it was found that this adsorption procedure was relatively inefficient when compared with precipitation by ammonium sulfate and alcohol. When both the precipitation and the adsorption procedures were used in conjunction, there was no increase in the activity of the final preparation. The purified product contains about 16 per cent nitrogen, shows all the protein color tests, and is precipitated and denatured as is a typical protein. It was found that denaturation of the protein was proportional to the loss of enzymic activity. The authors believe that their purified material appears to correspond to the  $\beta$ -amylase discussed by Kuhn (34).

#### CHEMICAL NATURE OF CARBONIC ANHYDRASE

In 1932, Roughton and Meldrum isolated from ox blood a white substance of which 1 part in 10,000,000 is able to catalyze the formation of carbon dioxide from a buffered bicarbonate solution of a physiological pH. The dry substance appears to be free from hemoglobin and heme compounds. It has the typical characteristics of an enzyme, but differs from any of the known enzymes of the blood. As to the chemical nature of the purest preparation, it resembles that of a protein and it gives all the protein color tests. Its molecule, however, seems to be of a small size judged by the fact that the enzyme tends to pass through ultrafiltration membranes. The physiological function of the carbonic anhydrase is to accelerate the carbon dioxide formation from bicarbonate in the pulmonary blood vessels. The possibility of numerous other functions within the animal body is suggested, such as the removal of carbon dioxide from active tissue, catalysis of decarboxylation, and the formation of urea in the liver. An account of the discovery of the new enzyme, purification, activity determinations, etc., and review of the literature has been given in an article by Roughton (65).

## THE CHEMICAL NATURE OF CATALASE

Catalase is found in abundance in animal and plant tissues. It decomposes hydrogen peroxide into water and oxygen, but it does not decompose other peroxides. Physiologically it is very important, since it destroys hydrogen peroxide (a toxic intermediate product of oxidation) within the living tissue. Another biological function of catalase is its dehydrogenation function. It furnishes oxygen for dehydrogenation.

Michaelis and Pechstein (51) and Stern (68) describe the enzyme catalase as an ampholyte of a high molecular weight, with an isoelectric point of about pH 5.5. Waentig (88) found catalase to be a protein containing 14 per cent nitrogen. Stern (69) determined the diffusion velocity of the catalase particle and states that it is similar to that of hemoglobin. According to Zeile and Hellstroem (100), catalase is a heme (ironporphyrin) compound. Kuhn, Hand, and Florkin (35) claim peroxidase to be a heme compound. This could not be corroborated, however, by Elliot and Sutter (11). Boyd (5) confirmed Howell's (23) earlier finding that hematoporphyrin in the presence of light hydrolyzes fibrinogen, so that it loses its power to clot with thrombin or coagulate on heating; nor does it precipitate upon half-saturation with sodium chloride. Furthermore, he was able to digest serum albumin in the presence of light with hematoporphyrin. Oxygen is essential for hydrolysis. "Hematoporphyrin unites with the protein, becomes excited through the absorption of light energy, combines with oxygen, and passes a part of the energy over to the protein molecule, thereby activating the latter and causing it to become hydrolyzed." Boyd calls hematoporphyrin an artificial proteolytic enzyme.

## DIGESTION OF ENZYMES BY PROTEASES

As shown above, the carrier theory, based largely on theoretical considerations, has been found to be untenable. Obviously, some method must be employed to throw light on this problem, and the digestibility of enzymes by proteases has recently been used to a considerable extent.

Tauber (77) found crystalline urease to be digestible by trypsin in the presence of a colloid such as gum arabic. The time required for complete inactivation of the urease was three days. Tauber and Kleiner (80) found gum arabic or gum ghatti to be essential for the digestion of urease by trypsin and found that no digestion of the urease took place if the addition of the gum was omitted, thus confirming the finding of Waldschmidt-Leitz and Steigerwald (92), that urease cannot be digested by trypsin in the absence of a gum. If trypsin is added to urease before the addition of the gum, no inactivation takes place. An inhibitive (antitryptic) group of the urease molecule combines with trypsin to form an irreversible inactive

trypsin compound which cannot digest the urease but has unchanged ureolytic activity. In the presence of a gum, this compound cannot form because the gum unites with a part of the urease molecule. Tryptic activity is unimpaired and the urease is digested. Tauber and Kleiner suggested that crystalline urease may possess "antitryptic" properties. Woodhouse (99) reported that gum arabic and gum tragacanth were able to augment definitely the hydrolysis of olive oil by lipase. According to Sumner, Kirk, and Howell (72) urease is rapidly inactivated by pepsin and by papain-hydrogen sulfide, and inactivation coincides with the digestion of the protein crystals.

It has also been shown by Tauber and Kleiner (82) that maltase is readily digested by trypsin. Tauber (78) found emulsin to be indigestible by pepsin, trypsin, and pancreatin during several weeks incubation at 40°C. Protein-free emulsin, however, could not be obtained by methods of earlier investigators, nor could the active principle be freed from the protein by other procedures. At the present time, no explanation can be given for the indigestibility of emulsin. It should be noted, however, that not all proteins are digestible by trypsin and pepsin, and that emulsin may not be a protein of the ordinary type. As an example, the heminprotease of Haurowitz (18) may be mentioned. Haurowitz digested oxyhemoglobin, CO-hemoglobin, and reduced hemoglobin by pancreatic extracts, all at approximately the same rate. Putrefactive bacteria have an inhibitive effect on the digestion. About 95 per cent of the globin is split off as dialyzable products and the residual 5 per cent forms a new combination with the protohemin fraction, which is amorphous and which Haurowitz calls heminprotease. This protein component is resistant to pepsin and trypsin. Rennin, as found by Tauber and Kleiner (81), is digested by trypsin with amazing rapidity. Pepsin digests rennin also with a fair degree of rapidity, but erepsin does not. The digestion of rennin by pepsin supports earlier evidence which Tauber and Kleiner (81) brought forward for the protein nature of rennin, and for the identity of rennin as an entity distinct from pepsin. Therefore, the claim that rennin exists in the adult mammal's stomach is quite paradoxical. Moreover, it has been experimentally proven that there is no rennin in the adult mammal's stomach (83). Another point of differentiation between these two enzymes is the difference in the pH of activation of their respective zymogens. Recent experiments of Dr. I. S. Kleiner and the author, to be published soon, corroborate their earlier findings as to the existence of individually differing zymogens of rennin and pepsin.

Schulman and Rideal (66) studied enzyme inactivation and distinguished between two activities,—bulk phase reaction and surface reaction. Pancreatin solutions were digested by pepsin at pH 2, and the pH of the solu-

tion was adjusted to pH 8. For "bulk phase reaction" no tryptic activity was observed; monolayers ("surface reaction") of egg albumin, however, are readily digested. Trypsin solutions which have been boiled at pH 8.0 behave in a similar manner. The authors' conclusion is: "In the tryptic enzymes, the 'free enzymes' can be separated from their protein-like carriers, but when free can only react with proteins when presented to them in a suitable form such as a monolayer at an air-water interface." They found, however, that proteases inactivated by acids or alkalis or poisoned by heavy metals lose both bulk and surface activities equally.

Tauber and Kleiner (86) subjected pepsin and trypsin to reciprocal digestion. At pH 2.0 pepsin completely digests trypsin, leaving the former unchanged. This takes place in a relatively short time. At this pH the trypsin is inactive but not destroyed. If, however, a mixture of the two enzymes is kept at pH 5.5, which does not permit peptic activity (but does not inactivate pepsin), the pepsin is digested by trypsin. That trypsin may be digested by pepsin had also been found by Long and Johnson (44), by Long and Hull (43), and by Northrop and Kunitz (60).

Salivary amylase is slowly inactivated by proteases (86). Euler and Josephson (12) claimed that yeast invertase is a protein, offering important evidence for their assumption. Willstätter (96), however, found that this enzyme could be obtained free or almost free from protein, carbohydrates, and phosphorus. This preparation contained varying amounts of nitrogen which did not parallel the activity of the invertase. Willstätter, Kuhn, and Graser (97) explain the tryptic inactivation of invertase, as carried out by Euler and Josephson, as due to the digestion of the protein carrier. According to Willstätter and associates, as has been pointed out above, and contrary to Euler and others, enzymes are associated with carriers which have nothing to do with the catalytic effect of the enzyme per se.

We have seen that a great many enzymes have been found to be digested by proteases and some of them with amazing rapidity. Many of these enzymes have been reported to contain little nitrogen or else to be "free of protein." If this is the case, their inactivation by proteases must be attributed to digestion of their protein carriers. The question arises: Is the carrier, if it really exists, always, or almost always a protein? If this is the case it is contrary to the basic principle of the carrier theory, i.e., that the carrier may be *any* colloid of high molecular weight.

#### THE SPECIFICITY OF CARBOHYDRASES

The theory of Leibowitz (1925) (39, 40), states that there are two kinds of *maltases*. One, which hydrolyzes both maltose and  $\alpha$ -methylglucoside, and another, which attacks only maltose. The latter is found in certain moulds, the former in yeast. Fischer and Niebel (13) found, as early as

1896, that the maltase of horse serum would hydrolyze only maltose and not  $\alpha$ -methylglucoside. Weidenhagen (95), however, denies the existence of a specific sucrase, maltase, or  $\alpha$ -methylglucosidase. His theory is that sucrose is split by  $\alpha$ -*n*-glucosidase because it is an  $\alpha$ -*n*-glucoside and by  $\beta$ -*h*-fructosidase because it is also a  $\beta$ -*h*-fructoside; maltose, however, being only an  $\alpha$ -*n*-glucoside is hydrolyzed only by  $\alpha$ -*n*-glucosidase. Weidenhagen extends the same view to other saccharides. Much interest has been aroused in this new theory, but it has not been generally accepted.

For example, Karström (29) found that the enzyme of a certain strain of *Bacterium coli* hydrolyzes maltose but no sucrose, as would be expected according to Weidenhagen's theory. This was confirmed by Myrback (52) and by Tauber and Kleiner (82). Others, like Pringsheim, Borchard, and Loew (63), found malt extracts and certain moulds to be inert toward  $\alpha$ -methylglucoside but not toward maltose. Ivanoff, Dodonowa, and Tschastuchin (26) obtained from mushrooms a maltase which hydrolyzed maltose but not sucrose. Schubert<sup>2</sup> prepared in Nelson's laboratory an invertase from honey, which hydrolyzed maltose but not  $\alpha$ -methylglucoside (53). A maltase inactive to sucrose and  $\alpha$ -methylglucoside has been extracted by Kleiner and Tauber (31) from mammary tissue. Recent experiments of Grassmann and associates (16, 17), using various sugars and glucosides as substrates and working with different carbohydrases, point to the invalidity of Weidenhagen's theory.

More recently Tauber and Kleiner (87) have tested the theory of Weidenhagen once more. Since he has stated that  $\alpha$ -phenylglucoside is more readily hydrolyzed by maltase than  $\alpha$ -methylglucoside, Tauber and Kleiner employed the new substrate. They used the enzyme obtained from the tropical plant *Solanum Indicum*. This maltase did not hydrolyze  $\alpha$ -methylglucoside. It did, however, slowly split  $\alpha$ -phenylglucoside, and maltose was very rapidly hydrolyzed. Therefore, it was suggested that the maltases be divided into two groups. The first, or *true*  $\alpha$ -glucosidases, split all  $\alpha$ -glucosides and maltose, e.g., yeast maltase; the second, or *pseudo*  $\alpha$ -glucosidases, split maltose, are only slightly active to certain  $\alpha$ -glucosides, and are inactive to all others. To the latter group belong maltase of *B. coli*, maltase of the mammary gland, malt extracts, and the maltase of *Solanum Indicum*.

Regarding the *specificity of the saccharases*, there is no general accord. Kuhn and Rohdewald (36) maintain that there are two types of saccharases. Some taka-saccharase preparations do not hydrolyse raffinose, while others (37) do. Leibowitz and Mechlinski (40) repeated these experiments and found that some taka-saccharase preparations contain

<sup>2</sup> An extensive review on the chemistry of invertase has recently been published by Nelson (53).

melibiase, which is responsible for the hydrolysis of raffinose, while others do not. In the former instance, the trisaccharide was hydrolyzed by the melibiase to galactose and sucrose, and the sucrose freed in this manner was then split by the taka-saccharase. Weidenhagen (94) contradicts this. He found that such taka-saccharase preparations as contained no melibiase could attack raffinose.

Tauber and Kleiner (87) obtained an enzyme preparation from *Solanum Indicum* which readily split raffinose, but both melibiase and sucrase were present.

#### THE EFFECT OF HIGH PRESSURE ON ENZYMES

Macheboeuf, Basset, and Levy (47) studied the effect of very high pressures upon enzymes. At 6000 atmospheres enzymes are not destroyed; at 17,000 atmospheres (ultrapressure), however, varying destruction took place, which depended on the pH of the solution, the duration of exposure, and other factors. Enzymes in general are less resistant than bacterial spores, more resistant than non-spored forms, viruses, or bacteriophage, and approximately as resistant as bacterial toxins. Certain protein solutions were found to coagulate under similar conditions. It appears that the inactivation of certain enzymes by high pressure is due to the destruction of the colloid molecule, which in most cases is a protein.

#### ENZYMIC HYDROLYSIS OF THE DIKETOPIPERAZINE RING

All proteolytic enzymes split peptide linkages. This has been explained by the fact that the amount of amino groups formed is equal to the amount of carboxyl groups liberated. That this is the case in peptic hydrolysis has been definitely established recently (R. K. Cannan (1930)). However there are exceptions, and, contrary to the earlier general belief that the diketopiperazine ring cannot be split by proteases, Abderhalden and Schwab (1) succeeded in hydrolyzing leucylglycyltyrosine anhydride, leucylglycylleucine anhydride, and leucylglycylserine anhydride by erepsin and by trypsin, respectively. The above diketopiperazine compounds have been obtained in stable form for the first time. Matsui (49) has reported similar results using asparaglyldiglycyltyrosine, a synthetic tetrapeptide, as a substrate. Pepsin, trypsin, and erepsin did not split this compound, but trypsin hydrolyzed it very rapidly. The same author (50) found that synthetic diketopiperazinecarboxylic acid was hydrolyzed by trypsin and trypsin, but not by pepsin nor by erepsin. Ishiyama (25) extended these observations by studying a series of diketopiperazine compounds, such as glycylaspartic anhydride, glycylglutamic anhydride, aspartic anhydride, glycylglutamine anhydride, glycine anhydride, pyrrolidonecarboxylic acid, and pyrrolidonecarboxamide. The



first three compounds were not split by erepsin nor by pepsin, whereas trypsin and especially trypsin kinase were very active in hydrolyzing them. The remainder of these compounds were not hydrolyzed by any of the proteases or by the erepsin. It appears that in the diketopiperazine molecule a free carboxyl group is essential if it is to be hydrolyzed by trypsin; this was the case in the first three compounds. This corroborates earlier findings. Amidation of the carboxyl group is sufficient to make these compounds which would ordinarily be digested, indigestible; for example, glycylglutamine anhydride is very resistant (see also experiments of M. Bergmann).

In view of the increasing number of the various biochemical catalysts and to assure a distinct differentiation between enzymic and non-enzymic catalysts, the author proposes the following classification:

#### CLASSIFICATION OF BIOCHEMICAL CATALYSTS

1. *Specific, cell-independent, biochemical catalysts or enzymes*:—catalysts produced by the living cell, but whose action is independent of the living cell, and which are destroyed if their solutions are sufficiently heated. Examples: pepsin, trypsin, maltase, lipase.

2. *Specific, cell-dependent, biochemical catalysts*:—catalysts produced by the living cell, active in vitro as well as in vivo, their activity, however, depending on the unimpaired cell. They are destroyed on heating and their activity ceases on mechanical destruction of the cell. Example: the catalyst effecting the synthesis of urea in the liver. Discussed by Krebs (33).

3. *Non-specific biochemical catalysts*:—catalysts elaborated by the living cell, their action being independent of the life processes of the cell. They are not destroyed when their solutions are heated. Examples: glutathione, ascorbic acid.

Considering the existing evidence it may be said that, with a few possible exceptions, enzymes are proteins or broken-down products of proteins. As has been seen, there are enzymes which are simple proteins (pepsin, an albumin; urease, a globulin), and some are derived proteins (rennin, a thioprotease). The preparation of certain enzymes by prolonged autolysis and further purification, their subjection to proteolytic digestion, and the study of chemically active groups in the enzyme molecule may throw further light on the chemical nature of enzymes.

#### REFERENCES

- (1) ABDERHALDEN, E., AND SCHWAB, E.: *Z. physiol. Chem.* **61**, 212 (1932).
- (2) BAMANN, E., AND LAEVERENZ, P.: *Z. physiol. Chem.* **223**, 1 (1934).
- (3) BERSIN, TH.: *Z. physiol. Chem.* **222**, 177 (1933).

- (4) BERSIN, TH., AND LOGENMANN: *Z. physiol. Chem.* **220**, 109 (1933).
- (5) BOYD, M. J.: *J. Biol. Chem.* **103**, 249 (1933).
- (6) BRUCKE, E.: *Sitzber. Akad. Wiss. Wien Math. natur. Klasse Part 2*, 601 (1861).
- (7) CALDWELL, M. L., BOOHER, L. E., AND SHERMAN, H. C.: *Science* **74**, 37 (1931).
- (8) CARNOT, P., GERARD, F., AND MOISSONNIER, S.: *Ann. inst. Pasteur* **35**, 1 (1921).
- (9) DYCKERHOFF, H., AND TEWES, G.: *Z. physiol. Chem.* **215**, 93 (1933).
- (10) EDIE, E. S.: *Biochem. J.* **15**, 507 (1921).
- (11) ELLIOTT, K., AND SUTTER, H.: *Z. physiol. Chem.* **205**, 47 (1932).
- (12) EULER, H. VON, AND JOSEPHSON, K.: *Ber.* **57**, 859 (1924).
- (13) FISCHER, E., AND NIEBEL, W.: *Sitzber. preuss. Akad. Wiss. Physik. math. Klasse 5*, 73 (1896).
- (14) GLICK, D., AND KING, C. G.: *J. Am. Chem. Soc.* **55**, 2445 (1933).
- (15) GODDARD, V. R., AND MENDEL, L. B.: *J. Biol. Chem.* **82**, 447 (1929).
- (16) GRASSMAN, W., STADLER, R., AND BENDER, R.: *Ann.* **502**, 20 (1933).
- (17) GRASSMAN, W., ZECHMEISTER, L., TOTH, G., AND STADLER, R.: *Ann.* **503**, 167 (1933).
- (18) HAUROWITZ, F.: *Z. physiol. Chem.* **183**, 161 (1930).
- (19) HELLERMAN, L., PERKINS, M. E., AND CLARK, W. MANSFIELD: *Proc. Nat. Acad. Sci.* **19**, 855 (1933).
- (20) HOLTER, H.: *Z. physiol. Chem.* **196**, 1 (1931).
- (21) HOLTER, H.: *Biochem. Z.* **255**, 160 (1932).
- (22) HOTCHKISS, M., AND TAUBER, H.: *J. Immunol.* **21**, 287 (1931).
- (23) HOWELL, W. H.: *Arch. intern. physiol.* **18**, 269 (1921).
- (24) HOWELL, S. F.: *Proc. Soc. Exptl. Biol. Med.* **29**, 759 (1932).
- (25) ISHIYAMA, T.: *J. Biochem. Japan* **17**, 285 (1933).
- (26) IVANOFF, N., DODONOWA, E. W., AND TSCHASTUCHIN, W. J.: *Fermentforschung* **11**, 433 (1930).
- (27) JACOBY, M.: *Biochem. Z.* **262**, 181 (1933).
- (28) KARRER, P., AND ZEHENDER, F.: *Helv. Chim. Acta* **16**, 701 (1933).
- (29) KARSTRÖM, H.: *Biochem. Z.* **231**, 399 (1930-31).
- (30) KIRK, J. S., AND SUMNER, J. B.: *J. Biol. Chem.* **94**, 21 (1931).
- (31) KLEINER, I. S., AND TAUBER, H.: *J. Biol. Chem.* **99**, 241 (1932).
- (32) KLEINER, I. S., AND TAUBER, H.: *J. Biol. Chem.* **104**, 267 (1934).
- (33) KREBS, H. A.: *Ergebnisse Enzymforschung* **3**, 247 (1934).
- (34) KUHN, R.: *Ann.* **443**, 1 (1925).
- (35) KUHN, R., HAND, D. B., AND FLORKIN, M.: *Z. physiol. Chem.* **205**, 255 (1931).
- (36) KUHN, R., AND ROHDEWALD, M.: *Z. physiol. Chem.* **154**, 64 (1925).
- (37) KUHN, R., AND ROHDEWALD, M.: *Z. physiol. Chem.* **163**, 17 (1927).
- (38) KUNITZ, M., AND NORTHEROP, J. H.: *Science* **78**, 558 (1933).
- (39) LEIBOWITZ, J.: *Z. physiol. Chem.* **149**, 184 (1925).
- (40) LEIBOWITZ, J., AND MECHLINSKI, P.: *Z. physiol. Chem.* **154**, 64 (1925).
- (41) LEVENE, P. A., AND HELBERGER, J. H.: *Science* **73**, 494 (1931).
- (42) LÖVGREN, S.: *Biochem. Z.* **119**, 215 (1921).
- (43) LONG, J. H., AND HULL, M.: *J. Am. Chem. Soc.* **38**, 1620 (1916).
- (44) LONG, J. H., AND JOHNSON, A.: *J. Am. Chem. Soc.* **35**, 1188 (1913).
- (45) LOUGHLIN, W. J.: *Biochem. J.* **27**, 1779 (1933).
- (46) LUERS, H., AND BADER, J.: *Biochem. Z.* **190**, 122 (1927).
- (47) MACHEBOEUF, M. A., BASSET, J., AND LEVY, G.: *Ann. physiol. physicochim. biol.* **9**, 713 (1933).
- (48) MATHEWS, A. P., AND GLENN, J. H.: *J. Biol. Chem.* **9**, 29 (1911).

- (49) MATSUI, J.: J. Biochem. **17**, 163 (1933).
- (50) MATSUI, J.: J. Biochem. **17**, 253 (1933).
- (51) MICHAELIS, L., AND PECHSTEIN: Biochem. Z. **53**, 320 (1913).
- (52) MYRBACK, K.: Z. physiol. Chem. **205**, 248 (1932).
- (53) NELSON, M. H.: Chem. Rev. **12**, 1 (1933).
- (54) NORTHROP, J. H.: J. Gen. Physiol. **13**, 739 (1930).
- (55) NORTHROP, J. H.: J. Gen. Physiol. **15**, 29 (1931).
- (56) NORTHROP, J. H.: J. Gen. Physiol. **16**, 323, 339 (1932).
- (57) NORTHROP, J. H.: J. Gen. Physiol. **16**, 615 (1933).
- (58) NORTHROP, J. H.: J. Gen. Physiol. **17**, 165 (1933).
- (59) NORTHROP, J. H., AND KUNITZ, M.: J. Gen. Physiol. **16**, 261, 295, 313 (1932).
- (60) NORTHROP, J. H., AND KUNITZ, M.: J. Gen. Physiol. **16**, 267 (1932).
- (61) PEKELHARING, C. A.: Z. physiol. Chem. **35**, 8 (1902).
- (62) PERRIN, J.: J. chim. phys. **3**, 102.
- (63) PRINGSHEIM, H., BORCHARD, H., AND LOEW, F.: Z. physiol. Chem. **202**, 23 (1931).
- (64) PURR, A.: Biochem. J. **27**, 1703 (1933).
- (65) ROUGHTON, F. J.: Ergebnisse Enzymforschung **3**, 289 (1934).
- (66) SCHULMAN, J. H., AND RIDEAL, E. K.: Biochem. J. **27**, 1581 (1933).
- (67) SHERMAN, H. C., CALDWELL, M. L., AND DOEBBELING, S. E.: J. Biol. Chem. **104**, 501 (1934).
- (68) STERN, K. G.: Z. physiol. Chem. **204**, 259 (1932).
- (69) STERN, K. G.: Z. physiol. Chem. **217**, 207 (1932).
- (70) SUMNER, J. B.: J. Biol. Chem. **69**, 435; **70**, 97 (1926).
- (71) SUMNER, J. B.: Proc. Soc. Exptl. Biol. Med. **31**, 204 (1933).
- (72) SUMNER, J. B., AND HAND, D. B.: J. Am. Chem. Soc. **51**, 1255 (1929).
- (73) SUMNER, J. B., AND KIRK, J. S.: Z. physiol. Chem. **205**, 219 (1932).
- (74) SUMNER, J. B., KIRK, J. S., AND HOWELL, S. F.: J. Biol. Chem. **98**, 543 (1932).
- (75) SUMNER, J. B., LLOYD, O., AND POLAND, L. O.: Proc. Soc. Exptl. Biol. Med. **30**, 553 (1933).
- (76) SUNDBERG, C.: Z. physiol. Chem. **9**, 319 (1885).
- (77) TAUBER, H.: J. Biol. Chem. **87**, 625 (1930).
- (78) TAUBER, H.: J. Biol. Chem. **99**, 257 (1932).
- (79) TAUBER, H., AND KLEINER, I. S.: J. Biol. Chem. **92**, 177 (1931).
- (80) TAUBER, H., AND KLEINER, I. S.: J. Gen. Physiol. **15**, 155 (1931).
- (81) TAUBER, H., AND KLEINER, I. S.: J. Biol. Chem. **96**, 745 (1932).
- (82) TAUBER, H., AND KLEINER, I. S.: J. Gen. Physiol. **16**, 767 (1933).
- (83) TAUBER, H., AND KLEINER, I. S.: Z. physiol. Chem. **220**, 205 (1933).
- (84) TAUBER, H., AND KLEINER, I. S.: J. Biol. Chem. **104**, 259 (1934).
- (85) TAUBER, H., AND KLEINER, I. S.: J. Biol. Chem. **104**, 271 (1934).
- (86) TAUBER, H., AND KLEINER, I. S.: J. Biol. Chem. **105**, 411 (1934).
- (87) TAUBER, H., AND KLEINER, I. S.: J. Biol. Chem. In press.
- (88) WAENTIG, P.: Fermentforschung **1**, 165 (1916).
- (89) WALDSCHMIDT-LEITZ, E.: Z. physiol. Chem. **214**, 75 (1933).
- (90) WALDSCHMIDT-LEITZ, E., AND KOFRANYI, E.: Naturwissenschaften **21**, 206 (1933).
- (91) WALDSCHMIDT-LEITZ, E., AND REICHEL, M.: Z. physiol. Chem. **204**, 197 (1932).
- (92) WALDSCHMIDT-LEITZ, E., AND STEIGERWALD, F.: Z. physiol. Chem. **195**, 260 (1931); **206**, 193 (1932).

- (93) WARBURG, O., AND CHRISTIAN, W.: *Naturwissenschaften* **20**, 980 (1932); *Biochem. Z.* **254**, 438 (1932).
- (94) WEIDENHAGEN, R.: *Z. Ver. deut. Zucker-Ind.* **78**, 125, 242, 406 (1928).
- (95) WEIDENHAGEN, R.: *Ergebnisse Enzymforschung* **1**, 168 (1932).
- (96) WILLSTÄTTER, R.: *Ber.* **59**, 1591 (1926).
- (97) WILLSTÄTTER, R., KUHN, R., AND GRASER, J.: *Z. physiol. Chem.* **123**, 1, 45, 59 (1922).
- (98) WILLSTÄTTER, R., AND ROHDEWALD, M.: *Z. physiol. Chem.* **108**, 258 (1932).
- (99) WOODHOUSE, D. L.: *Biochem. J.* **26**, 1512 (1932).
- (100) ZEILE, K., AND HELLSTROEM, H.: *Z. physiol. Chem.* **192**, 171 (1930).



## COLLOID DEVELOPMENTS IN SYNTHETIC RESINS

H. L. BENDER, H. F. WAKEFIELD, AND H. A. HOFFMAN

*Research Division, Bakelite Corporation, Bloomfield, New Jersey*

*Received March 22, 1934*

Industrial resins are essentially organic glass-like mixtures which melt and flow at sufficiently low temperatures to enable reinforcing with a great variety of materials. They are alike in having great strength when in thin sections, that is, when approaching molecular thickness of films, but differ progressively from some resinous mixtures having fair strength in massive form to others having comparatively low massive strength. They may be brittle, may lack shock resistance, may be of low tensile strength as a mass, and yet may exhibit great strength when used as an adhesive between two surfaces, for instance, between two particles of wood. The essential factor then of the resinous field is the great increase in strength as the plastic flows and hardens in sections of near-colloid thickness.

This resinous plastic field may well be divided into two main divisions dependent upon properties which find reflection in use. (1) The *resins*, which are melted for flow and cooled for hardening into the finished shape. These are generally called thermoplastic resins and include natural resins, the resin esters or ester gums, the cellulose esters and ethers, and a great many synthetic resins, as the novolak phenol-aldehyde resins, the glyptal resins, the vinylite resins, etc. (2) The *resinoids*, which in molding are heated for flow and also heated for hardening effect. These include the original heat-hardening phenol-formaldehyde resinoids of Baekeland and their many modifications, the urea-aldehyde resinoids, and some of the glyptal type.

The concepts of film formation, adhesion, surface tension, etc., are perhaps common to the above two groups; the two groups differ sharply in permanency of plasticity and in the importance of particle size effect on properties.

Resinous compounds and their mixtures differ from crystalline compounds in the capacity for plastic flow at temperatures below a true melted condition as contrasted to the sudden transition from a rigid condition to the limp liquid flow of melted crystals. This property of plastic flow with heat naturally leads to analogy with the fields of metal, glass, and rubber. These fields, being older and easier to study, have to date given us the more valuable colloid data bearing on the resin field. For example,

present empirical resin formulas take advantage of the ability of resins to temper like glass and metals, and to stiffen and strengthen like rubber with the incorporation of small particles. Yet such general data need better assembling and classification so that they can find their true sphere of use in the resin industry.

Data for the resin field might well be divided into three groups: (1) Flow, or plasticity, which deals with the resin mixtures in motion and with resistance to flow. (2) Rigidity, set, or hardness, which deals with a static condition or total resistance to flow. (3) Colloidal surface effects, or distensibility both of portions within the mass and of the mass as a whole distributed over added particles.

Plastic flow is a universal property of matter and dependent only upon the pressure exceeding a definite yield value. Thus the familiar collapsible metal tube for tooth paste may be cold-pressed to shape and even steel cold-drawn into seamless tubing. Reiner (38) says that every material flows under certain conditions and refers to Karman (24), who described cubes of marble flowing under sufficiently high pressure as behaving under those conditions like any plastic substance. The problem in plastic materials then is not just to make materials flow but to control the flow of certain adhesive materials under available commercial pressures.

We see that plastic flow involves mechanical work, so we are induced to consider applied force in different ways: first, as the cause of controlled plastic flow; second, as the cause of changes in plastic flow values; and third, as will be explained later, as a cause of certain definite changes in physical properties.

In the resin plastic field properly controlled plastic flow is obtainable as follows: (1) by choice of resinous material; (2) by choice of filling materials, both coarse and fine; (3) by fineness of grinding of the solid mixed mass with a proper balance of fine and coarse particles; (4) by processing the resin to proper plastic flow range, as by means of polymerization, by distilling off liquid products, or by addition of plasticizers; and (5) by the use of solvents or temporary plasticizers.

For the use of coarse filling materials we have Baekeland's (3) original rule: "Judicious incorporation of fibrous substances provides increased resistance to shock in the final heat-hardened product."

For the use of colloidal filling materials we find data available in the rubber field. Rubber, under proper conditions, fulfills all the essentials of resins. Except for the fact that rubber has grown into an immense special flexible field of its own, the resins and rubber fields might be classed as a unit.

We shall consider later the definite resinous properties of rubber, but we are here concerned with the availability of data concerning additions of

small-sized particles. Endres (14) has shown the effect of particle size of barium sulfate from colloid size up to gross size. He found an increased stiffening and strengthening effect with decrease in average size. Apparently the effect reaches a maximum such as Alexander's (1) "zone of maximum colloidity," since this effect is known to disappear again when we reach the realm of dispersed molecules.

Polymerization in a mass at rest is often a very different phenomenon from that of the same mass when in motion. Polymerization is the term we use to denote change of a substance without loss or gain of material but generally, with a transfer of energy, to a less soluble, less fusible, less chemically active form of higher average molecular weight. We have many examples both of easily reversible and of so-called irreversible polymerization. Naturally the most complete data has been published in the field of reversible polymers.

Staudinger (46) has many papers on the polymers of styrene, some forms of which are controllable and reversible. He may be considered the foremost exponent of a micellar structure as formed from primary valences, just as Meyer (35) expounds the merits of the secondary valence or association theory, and Bergmann the theory of coordination in the solid state of units incapable of free existence. All the various theories of polymerization are intensely interesting, but the known facts are even more interesting.

Any simultaneous lowering of flow, lowering of solubility, raising of melting point, and raising of average molecular weight is by definition polymerization. We at present suspect that this broad definition may cover two distinct classes of phenomena: (1) Change in particle or molecular size, generally resulting in higher average size but not necessarily higher maximum size. (2) Rearrangement as to particle or molecular position, that is, a change in average directional influence.

Staudinger (47) has given us some very valuable suggestions as to long slender chains of monomeric groups linked together by primary valence bonds. Length of the chain forms his distinction between hemicolloids and encolloids. The short chains of a hundred molecules or so are said to be stable to heat treatment when in solution, whereas the longer (and reversible) encolloids of some thousand molecules often lose viscosity by heat treatment. Solutions of these long fiber molecules form films which are like mats of macrofibers and are tough and strong from their interlocking structure. Thus from styrol at elevated temperatures, hard, brittle masses were obtained of comparatively low molecular weight, while from long-time reaction at low temperatures, tough masses were obtained which were more flexible and had average molecular weights four or five times as large.



We have found that the picture as drawn by Staudinger from reversible polymers cannot be indiscriminately used in the heat-hardening field. At any temperature, solutions of resinous particles of the heat-hardening type show the effect of conflicting forces, various ones of which predominate at different times and under different conditions. When one suspends solid resinous particles in a solvent, one expects a change in viscosity until a stable condition is reached. In the case of synthetic resinoids one notices a swelling of the particles, followed by dispersion with a drop in viscosity. If a condensation or polymerization reaction is progressing at the temperature of test, one then may expect a gradual rise of viscosity with time, the measured viscosity at any instant being, of course, the resultant of those several forces. When the resinoid is heat-hardened in the presence of the solvent, the swelling and dispersion phenomena may be eliminated from consideration and we should expect a gradual rise of viscosity. This is a beautiful theoretical picture, but the fact remains that almost all such resinoid solutions show a sudden and sharp viscosity drop, which may be completed a few minutes or hours after cooling to room temperature.

Since these changes are rapid and nearly always complete by the time a still charge is ready for shipment, they were ignored until by accident we found formulas in which the changes were prolonged sufficiently to interfere with commercial use. In figure 1, curves A and B are the changes in viscosity with age of two commercial rapid heat-hardening phenolic resinoid varnishes. Curve D shows changes of an experimental slow type flexible phenolic resinoid in one solvent, and curve C shows the changes of the same resinoid as in D, but in a mixture of solvents.

In this class of resinoid solutions we find, as also was found by Carothers and Hill (7) in the resinous ester field, and contrary to Staudinger's theoretical concepts, that high temperature favors high viscosity and low temperature favors low viscosity. Perhaps as yet no general rule can be drawn. We had the case of a solution which having been heated to 3000 centipoises at 95°C., then cooled to 25°C., dropped in four weeks to 500 centipoises. At the end of a year it tested 600 centipoises. The next surprise was to find resinoid solutions showing unstable viscosity at constant temperature, the fluctuations occurring between non-flowing gels and low viscosity liquids over constant and reproducible cycles as shown by the curves in figure 2.

Some changes in plastic flow phenomena occur in raw rubber when it is "racked," that is, pulled into a thread at any temperature and held rigid by excessive cooling under stress (17). On again being heated the rubber will retract at a temperature which bears a definite relation to the temperature at which stretching occurred. Bary and Hauser say that rubber

consists of two polymers,  $\alpha$ -rubber and  $\beta$ -rubber, the  $\beta$ -rubber being the less soluble. X-ray examination of "racked" rubber shows that the

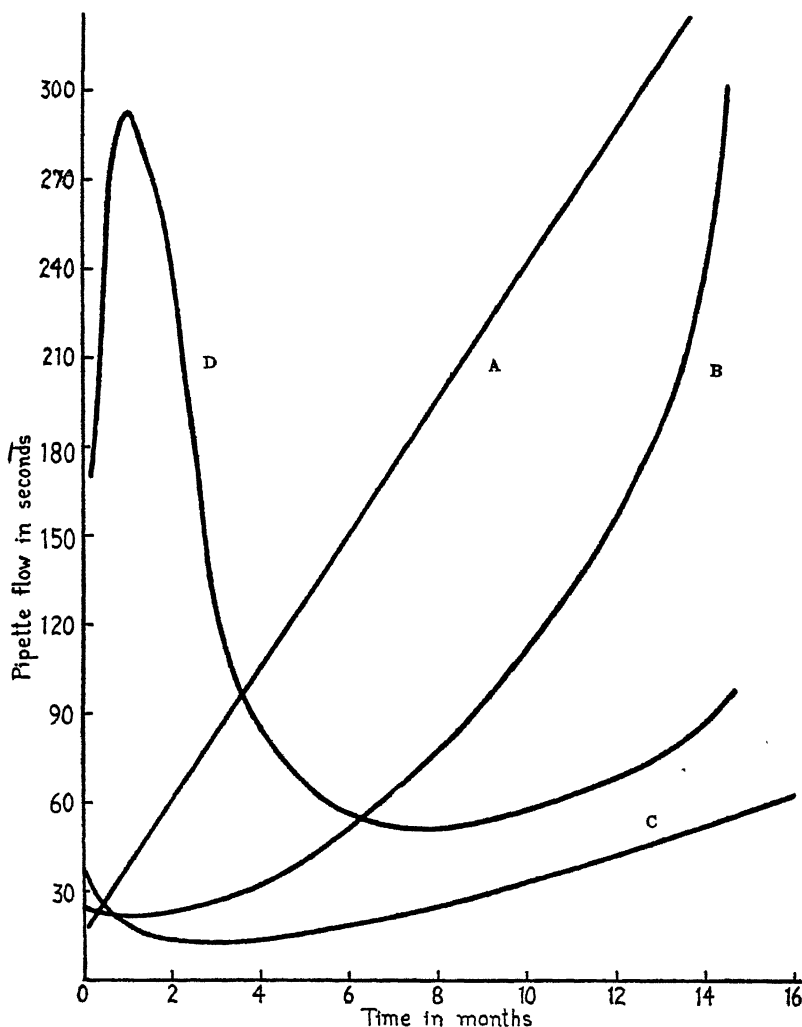


FIG. 1. VISCOSITY CHANGES IN SOLUTIONS OF RESINOIDS STANDING AT ROOM TEMPERATURE

A, commercial phenolic-formaldehyde resinoid varnish; B, rapid hardening phenolic resinoid solution; C, flexible mixed phenols-formaldehyde varnish in mixed solvents; D, same resinoid as in C, but in a single solvent.

change from  $\alpha$ - to  $\beta$ -rubber can be followed in the racking process, with the probability that the  $\beta$ -form is in a more highly aggregated condition (4).

Since the racking of rubber turns it temporarily into a resin-like compound, the data become of great importance in the resin field.

The data by Gibbens, Gerke, and Tingey (17) shows that for racking temperatures above 20°C. there is a fairly uniform difference between the

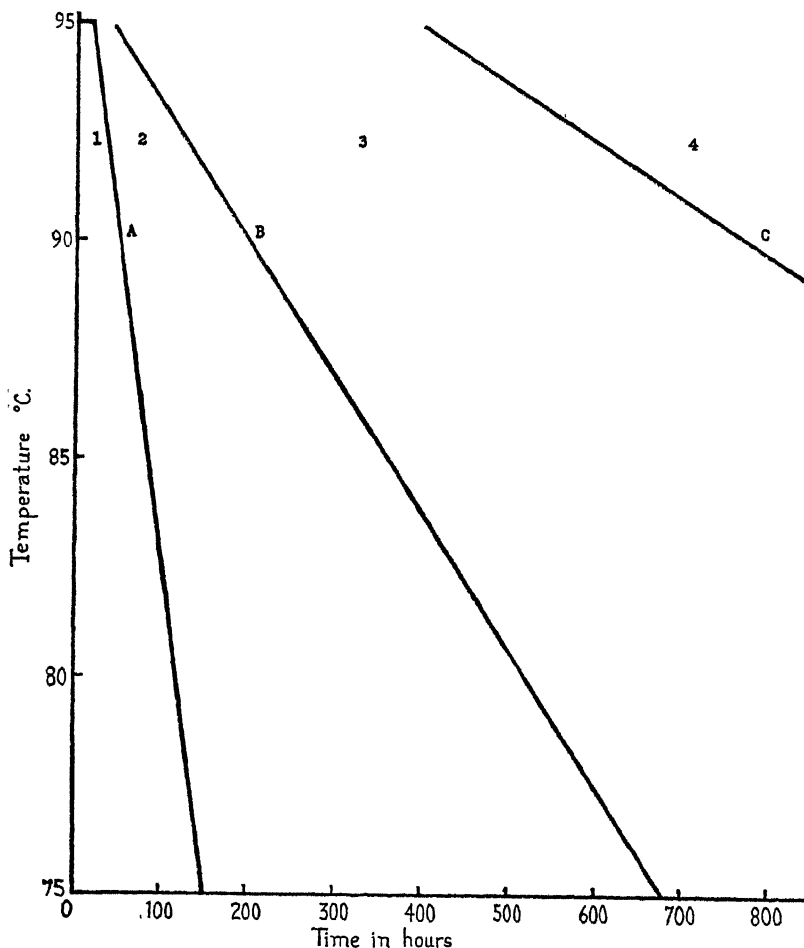


FIG. 2. GELATION CHANGES OF PHENOLIC RESINOID SOLUTIONS

A, time of gelatin of XH-1383 (60 per cent phenolic resinoid solution); B, time of reliquefaction of XH-1283 solution; C, time of regelation.

1, first liquid state; 2, first gel state; 3, second liquid state; 4, second gel state.

racking and the retraction temperatures. This temperature difference may be a measure of the molecular or micelle rearrangement resulting from the mechanical stretch.

A glance at the properties of "racked" rubber as compared to "unracked" rubber will indicate that we are here dealing with a reversible polymerization change induced by mechanical work. "Racked" rubber has increased in density (15), has taken on a fibrous structure (20), has increased in melting point (15), has increased in breaking strength, has become hard and inelastic, has given up heat, and has become of lower solubility.

$\alpha$ -Rubber and  $\beta$ -rubber are considered as different polymers of rubber hydrocarbon modified perhaps by natural proteins, resins, etc. In this connection it is interesting to note that pure rubber hydrocarbon when crystallized exists as fine needles (42). On the other hand, rubber goes to a resin-like product when heated in air with catalysts present (13).

All high-melting resins and some non-melting resins approach a rubber-like state at elevated temperatures. Rubber itself is a hard brittle body at very low temperatures. Rubber then might be considered as a fortuitous case of having elastic properties at room temperatures.

Autopolymerized styrene at temperatures above 65°C. behaves essentially as does raw rubber at room temperature. Thus above this temperature range the styrene polymer may be "racked." It then shows elastic set and elastic after-effects, as disclosed by Whitby and coworkers (50).

Another known effect of mechanical work on plastic flow may be found in the studies by von Weimarn (48) of natural silk dissolved and precipitated from solution. He has proposed the descriptive phrase "coagulation by motion."

von Weimarn dissolved (disintegrated) natural silk, then coagulated and dried the coagulant. Instead of a flexible substance like silk he arrived at a brittle resinous "glass-like" mass. When the wet mass was pulled into threads, using one simple pull, and the relaxed thread then dried, a brittle thread of low strength was obtained. However, when the thread was stretched and allowed to relax several times before drying, much stronger threads were obtained. From this discovery it was but a step to finding the mechanical working conditions which give silk threads a degree of strength far surpassing that of natural silk.

This concept of strengthening fibroin into silk thread is an exceedingly interesting one, especially when we note that the silk worm stretches and manipulates the secreted fibroin and that such stretching causes it to congeal rapidly.

To designate the concept of increased strength as derived from motion resulting from externally applied force we shall for convenience use the term "fibration." Thus natural silk is said to be "fibrated" fibroin. This concept of silk as a "fibrated" or rearranged resin will in time both explain many phenomena in the silk field and lead to better understanding in the resin field. Consideration of "racked" rubber and "fibrated" silk will

explain empirical data in the resin field, and perhaps systematic search will lead to many resins which can be "fibrated" to flexible threads or "un-racked" to rubbery masses.

Carothers and Hill (8) produced high molecular polymers by the reaction of dibasic acids and dihydric glycols. They assume a theoretical completeness of this reaction series when all the initial reactants have united, into one single massive molecule with a progressive loss of water. When these reaction products reach an apparent molecular weight of around 9000, they are generally resinous in properties. Certain of these high molecular weight products exhibit the peculiar property of a temporary plastic flow which Carothers (9) has characterized as "their capacity to be drawn out into very strong, pliable, highly oriented fibers." That is, in the preferred terminology these products can be "fibrated." They disclose that a hot-drawn or extruded rod of these resinous materials is opaque, brittle, and resin-like, but if the same be subject to tension at lower temperatures, it will undergo a limited amount of flow and then, as an effect of this motion, will build up resistance to flow, that is, as we say in the plastic field, it will harden or congeal. These rods are therefore rendered less plastic by tension. By flow they have been transformed from an opaque to a transparent mass of high luster. They have increased greatly (about 600 per cent) in breaking strength and in pliability. The final strengths obtained were about equal to the values for cotton fiber.

Von Weimarn found that the rearrangement in fibroin could be followed under the microscope. The "fibration" proceeds from the outside towards the center with increased mechanical manipulation.

Some changes of properties with mechanical work are well-known for many compounds and mixtures. For instance, ordinary gelatin-water mixtures decrease in viscosity on stirring, so much so that in ice cream laboratory technique the "basic viscosity," defined as the minimum obtained on stirring, is reported (36). The viscosity of resin solutions and certain properties of solid resins, such as melting points, are also subject to the same type change with mechanical agitation, and this fact must often be considered to avoid confusion of data.

In weak gel-like solutions of these amorphous materials we note two conflicting influences. (a) That which is commonly called "thixotropy" (25; also 43) is the gel property of being easily broken with agitation and of reassembling on standing. (b) Rearrangement to a flexible condition within the gel. This last is of considerable interest and is not so well-known.

A dilute ammonium oleate solution is a weak gel when standing; Seifriz (39) prepared a solution with a viscosity to just support a  $10\mu$  nickel particle. With mechanical agitation this solution assumes a very highly

elastic state. On cessation of rotary stirring, air bubbles may be withdrawn backward as much as  $180^\circ$ . When quiet and measured immediately, it shows a viscosity to support a nickel particle forty-five times as large as before agitation. This elasticity, however, is gradually lost on standing and finally reaches the former quiet value of 10. The process may be repeated many times with the same solution. Threads of this ammonium oleate gel may be stretched several inches.

In addition to the disruption of gels and the temporary strengthening of gels we also note that sols have been coagulated by stirring (16, 37).

Laszlo (28) considers that shellac, distyrene, and rubber are all two-phase gels, and that even hard glass-like resins may be jellies built up from and containing both liquid and solid constituents. If the hard resins are jellies we should then by analogy also find in them all of the above effects of mechanical work as are known to change the properties of gels, namely: (1) change of plastic flow, both loss and gain, as the result of motion; and (2) change of other properties as the result of motion.

The curves of various resinoids in figure 3 show the breaking-down effect of the resin structure under milling conditions as being typical for gel structures.

It is well-known that mastication of rubber directly influences the flow and vastly affects the viscosity of sols prepared from it (51). It is perhaps not so well-known that on mastication rubber at first becomes firmer and harder, and that this change soon gives way to an increasing softness, or that on allowing moderately worked rubber to stand it regains some of its firmness and elasticity (10).

It is also well-known that cold-worked metal often increases in stiffness, hardness, and brittleness, but it is not so generally understood that other properties also change. For instance, according to Austin and Pierce (2) the coefficient of expansion decreases as the result of cold-working.

In deforming crystals (26) the limit of elasticity can be raised several times by successive deformations. This strengthening in one direction is accompanied by weakening in another direction. Strengthened or weakened crystals after a time regain their normal condition.

Certain types of artificial stone are caused to set or harden by mechanical influences (29), and in clay articles we find differences between those made by casting and those made by extrusion.

Orientation in natural products has long been recognized. The wide occurrence of this phenomenon, as touched upon, for instance, by Bragg (5), seems to us to have a very special significance in surface phenomena. Hardy (18) early suggested that all molecules are more or less definitely oriented at interfaces, and this concept has been extended by Langmuir (27) for solutions.

Wilson (52) has proposed that pure liquids when drawn into thin films take on the properties of plastic solids, are stronger than the original liquid, may be stretched, and thus have elastic properties.

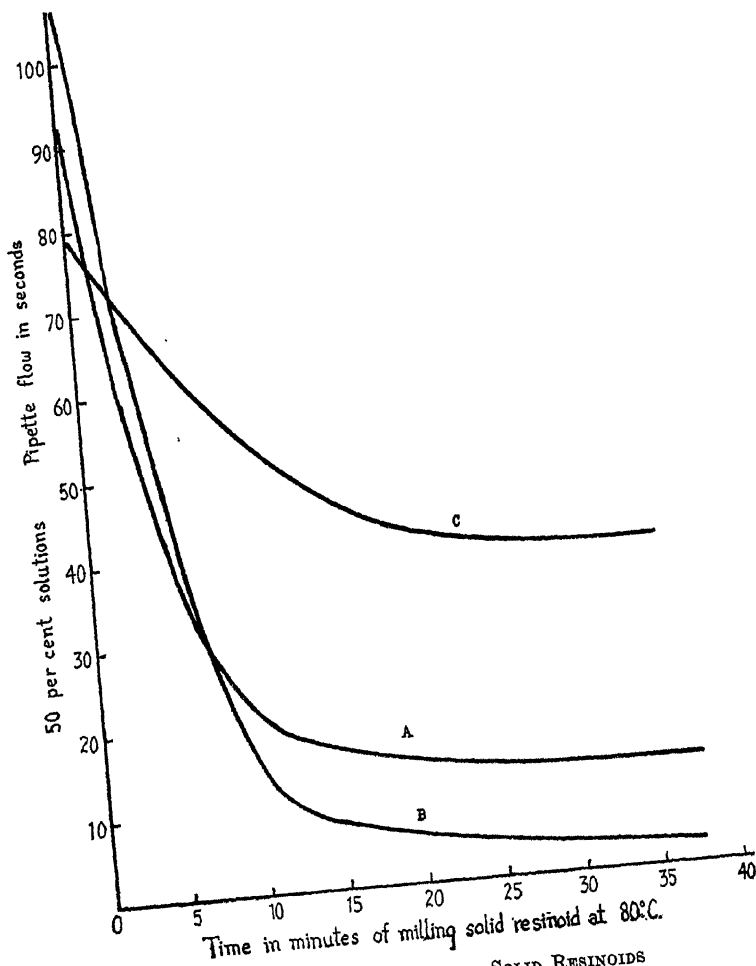


FIG. 3. EFFECT OF MILLING ON SOLID RESINOIDS

A, phenolic resinoid; B, alkyd resinoid; C, urea resinoid. Measured in solvents

The colloid chemists have long agreed on the great strength of the surface layer as contrasted with the lesser strength of the mass of the same material (34). It seems but a logical step to proceed from the consideration

of actual surfaces to consideration of films or surface skins of small but definite thickness.

If the effect of strengthening by flow is a true thin film effect as distinguished from a massive effect, one might find valuable data in the realm of films of molecular thickness. A recent publication (40) shows that thin film strengths may often increase with time, may reach high values, and may show remarkable elongation before breaking.

In "racked" rubber and in "fibrated" silk we appear to have film effects—either films in multiple layers or as threads—which may be considered as bundles of cylindrical films.

Clark (11) has given for stretched rubber a fairly complete picture closely agreeing with the above picture of stretched plastic silk. The two differ in one main detail, that is, the transformed plastic silk is permanent in form, as far as now known, whereas the racked rubber is permanent only below certain temperatures. For instance, if racked at 70°C. it is then permanent up to 40°C.

These "racking" and "fibration" or rearrangement effects in plastic substances have definite industrial value. This fact is becoming appreciated, as is shown by the current trend of research work. By orienting casein plastic a difference of strength of 600 per cent has been found in right angle directions. As the result of their studies, Manfred and Obrist (30) conclude that the flow of a casein plastic is lessened by the act of flow itself, whereas elasticity is increased by flow (31). Lately many patents have been issued on the stretching and working of spun rayon or cellulose derivatives. This operation gives increased strength (54), and involves a definite time factor for flow effects (55). The stretched fiber is far stronger to sudden load but is sensitive to increase of load per time unit, and under continued stress the strength fades to the same value as for unoriented rayon (22).

Sheppard and McNally (56) carried the flow effect into the film field, when for proper conditioning of the film they caused a film under tension to flow in a desired direction by means of pressure.

We now believe that many surface skin effects are caused by a type of reversible polymerization, as by rearrangement of micelles, and that such special surface films have different properties from the mass, that is, different average molecular weights, different solubilities, different strengths, densities, indices of refraction, etc.

The colloid concept of resinous micelles arranged in parallel layers, arrived at in one way by Staudinger and in another way by von Weimarn, will ultimately have a great influence on the study of resins. A recent diagrammatic representation of cellulose by Clark will be found of interest (12; also 41).



Staudinger found short isolated cylindrical bundles of fiber-like resinous masses, in form something like crystals but with amorphous properties. We note that slow long-time formation gave great micelle length, whereas short-time high temperature formation gave short lengths. This which we call skin effect looks like a directional influence analogous to the directional influences at work in crystal growth, but having perhaps a more profound effect on properties than is found in different crystal forms of the same substance. By intermittent longitudinal stretching and side compression von Weimarn was able to obtain progressively, from the surface inwardly, complete conversion of a mass of 0.02 mm. cross-section and of any convenient length, into a bundle of adhering threads.

Practical use of styrol resins came after control of resin quality by controlling the concentration of styrol in solution and its temperature of polymerization (49). An explanation of the varying degrees of toughness and brittleness is at hand in Staudinger's short and long micelles. Now a resin gel would seem to have a structural framework of long micelles with the physical nature and properties of the micelle influenced by temperature, by rate of formation, and by mechanical rearrangement.

This leads us to conceptions of resins as corresponding somewhat to that of cellulose. For instance, we can apply the cellulose concept of Meyer (35) that long primary valence chains unite to form a bundle (the micelle) through reciprocal lateral attraction. Also, Sponsler (44) interpreted the x-ray diagram of cellulose in terms of parallel oriented chains of unit groups held by primary valence.

Cellulose has properties which may some time greatly clarify the connection between structure and ultimate strength. Cotton fiber has tensile strength higher than that of copper (32). Regenerated cellulose is far less strong than the original cellulose, but it gains in strength when subjected to plastic flow such as stretching while in a swollen condition (23).

The natural cellulose structure, as in ramie cellulose, which consists of a series of concentric shells, shows great strength in one direction and much less strength at right angle thereto (45). Thermal expansion is ten times as great in one direction as in the right angle direction (19), and shrinkage is fifty to one hundred times as great transversely as longitudinally (6).

The cellulose fiber, for instance, cotton, is in the elongated or "fibrated" form of greatest strength, and any change is likely to be in the reverse direction, thus giving a resin-like and less strong massive form.

Mercerization or treatment with caustic soda solution causes the cotton fiber to shrink and weaken. It is therefore of great interest to find that caustic solutions have but little chemical effect below a certain concentration, but that such weak solutions have very great effect in causing contraction of the cotton thread (33). The contracted thread becomes

wiry and more brittle, that is more resin-like. The flexibility and strength of such thread may be partly restored by stretching the moist hot yarn. Mercerization of cotton increases iodine absorption, but when tension is used this iodine absorption is again reduced (21).

If in this resinous field needle-like amorphous micelles can form at the surface of plastic (supercooled melt) solids and can also form in solvents, then one should carefully consider cases where the melted plastic itself becomes the solvent. We have many cases of this in resins of abnormally high viscosity, high melting points, and abnormal solubility due, it would seem, to a polymeric structure within the mass. In the field of pure resinous masses one can find gels of every conceivable kind, reversible and irreversible, fusible and infusible, soluble and insoluble, hard and soft, etc.

According to Wright (53) alkyd resins are related in chemical and physical properties to drying oils. He says that gelation in both oils and alkyd resins depends upon escape of volatile products. In both cases degelling takes place if heated under pressure or in a confined space. Furthermore, while the alkyd resins may be plasticized to become an apparent physical match to celluloid, yet the celluloid mass may be molded and the alkyd structure will break down under the molding operation. Besides mechanical breakdown of the gel structure, the alkyd resin polymerization may be reversed by treatment with hot glycerol and also the resin may be vaporized, with decomposition, by application of dry heat above 250°C.

These polymerization phenomena are perhaps all reversible. For instance, hardened urea-formaldehyde resinoids dissolve easily in a 37 per cent water solution of formaldehyde, while phenolic resinoids dissolve with apparent depolymerization in phenolic solvents at about 300°C.

From this review of colloid literature we arrive at a common viewpoint of three fields, namely, rubbers, resins, and fibers; they seem to show some very close relationships. Rubber forms can be changed to resin forms and even to fiber forms. Resin forms can be changed to rubber forms and to fiber forms. Those products which are normally fibers may be changed to resins and to rubbers. Thus we are led to consider these molecules or micelles as mixtures of rods of varying size, kind, or condition, and to assign products to one or the other field depending upon rather general agreement in certain physical properties and irrespective of any definite agreement in chemical constitution.

With these similarities and differences in mind we may postulate a structural mechanism as a possible guide for further study of the connection between these three interlocking fields.

Having a mass with two or more rod forms present, let us assume no particular antagonism or attraction to exist between the different rods. Such a mass is an indiscriminate mixture with equal strengths in all direc-

tions. If, however, we push the pile, as into a thin plate, the motion will cause some rearrangement of the rods and increase in one direction the breaking strength of the oriented mass.

Now, if in a second case the rod ends are assumed to be points of mutual attraction, then as they slip past each other they should interact or interlock and we would have an endwise growth of micelles, or fiber-like forms. In case these new strong micelles resist chemical disruption, we would consider them as polymers. Rod forms of indifferent attractive forces would help to explain resins. Mutual attraction between rod forms would explain fibers and also explain some polymers, with all possible gradations in between.

Now having a mass with two or more rods present, let us assume a repulsive force between the ends of the rods. When such a mass is heated to a plastic flow range, the rods may be conceived to rearrange to greatest distance between ends, or at right angles to each other. Such a rearranged mass will resist and return from deformation, whether by stretching or by compression, and thus fall into the physical class of rubbers.

Rubber, resin, and fiber structures may thus be explained by the postulate that there are all possible variations of degree of attraction or of repulsion between rod forms of micelles. It should therefore be possible to pass temporarily from one state to another by simple motion, and, by assuming controlled variation in the degree of attraction, to explain those stable changes which range from minor to such major changes as transformation from one of these forms to either of the others.

The synthetic resin industry finds its major tasks in causing a controlled interruption, first, of chemical condensation, with simultaneous by-product elimination; second, of chemical polymerization, with the attendant change of particle size; and third, of physical polymerization, with corresponding change of properties due to enforced change of position. In the last two undertakings the principles of colloid chemistry find useful application.

#### REFERENCES

- (1) ALEXANDER: Colloid Chemistry, Vol. I, p. 20. The Chemical Catalog Co., New York (1926).
- (2) AUSTIN AND PIERCE: Ind. Eng. Chem. **25**, 779 (1933).
- (3) BAEKELAND: Ind. Eng. Chem. **1**, 157 (1909).
- (4) BARY AND HAUSER: Rev. gen. caoutchouc **42**, 3 (1928).
- (5) BRAGG: Colloid Chemistry **2**, 98 (1928).
- (6) BROWNE AND TRUAX: Colloid Symposium Monograph **4**, 263 (1926).
- (7) CAROTHERS AND HILL: J. Am. Chem. Soc. **54**, 1587 (1932).
- (8) CAROTHERS AND HILL: J. Am. Chem. Soc. **54**, 1559, 1566, 1569, 1579 (1932).
- (9) CAROTHERS AND HILL: J. Am. Chem. Soc. **54**, 1563 (1932).
- (10) CLARK: India Rubber World **79**, 54 (1929).
- (11) CLARK: India Rubber World **79**, 55 (1929).

- (12) CLARK: *Ind. Eng. Chem.* **22**, 474 (1930).
- (13) DAVIES: British patent 373, 228 (1932).
- (14) ENDRES: *Ind. Eng. Chem.* **16**, 1148 (1924).
- (15) FEUCHTER: *Kautschuk* **4**, 8, 28 (1928).
- (16) FREUNDLICH AND KROCH: *Naturwissenschaften* **14**, 1206 (1926).
- (17) GIBBONS, GERKE, AND TINGEY: *Ind. Eng. Chem., Anal. Ed.* **5**, 281 (1933).
- (18) HARDY: *Proc. Roy. Soc. London* **86A**, 610 (1912).
- (19) HENDERSHOT: *Science* **60**, 456 (1924).
- (20) HOCK: *Kolloid-Z.* **35**, 42 (1924).
- (21) HUBNER, J.: *J. Soc. Chem. Ind.* **27**, 105 (1908).
- (22) JASICEK: *Melliand Textilber.* **13**, 239 (1932); *C. A.* **27**, 3828.
- (23) JENKINS: *Chemistry and Industry* **52**, 242 (1933).
- (24) KARMAN: *Z. Ver. deut. Ing.* **55**, 1749 (1911).
- (25) KISTLER: *J. Phys. Chem.* **35**, 815 (1931).
- (26) KUZNETZOV AND SEMENTZOV: *J. Russ. Phys. Chem. Soc., Phys. Pt.*, **62**, 413 (1930).
- (27) LANGMUIR: *Chem. Met. Eng.* **15**, 467 (1916).
- (28) LASZLO: *Rubber Chem. Tech.* **2**, 576 (1929).
- (29) MANFRED AND OBRIST: *Plastics* **4**, 270 (1928).
- (30) MANFRED AND OBRIST: *Z. angew. Chem.* **41**, 971 (1928).
- (31) MANFRED AND OBRIST: *Plastics* **4**, 190 (1928).
- (32) MARK: *Trans. Faraday Soc.* **29**, 6 (1933).
- (33) MATTHEWS: *Chem. Engr.* **12**, 128 (1910).
- (34) MCBAIN: *Nature* **120**, 362 (1927).
- (35) MEYER, K. H.: *Biochem. Z.* **208**, 1 (1929).
- (36) MUELLER: *Ind. Eng. Chem.* **25**, 707 (1933).
- (37) OSTWALD: *Kolloid-Z.* **41**, 71 (1927).
- (38) REINER: *J. Rheology* **1**, 11 (1931).
- (39) SEIFRIZ: *Colloid Symposium Monograph* **3**, 293 (1925).
- (40) SERRALACH, JONES, AND OWEN: *Ind. Eng. Chem.* **25**, 819 (1933).
- (41) SHUTT AND MACK: *Ind. Eng. Chem.* **25**, 690 (1933).
- (42) SMITH, SAYLOR, AND WING: *Rubber Age* **33**, 162 (1928); *Bur. Standards, Paper No. 544*.
- (43) SNEDBERG: *Repts. Faraday Soc. and Phys. Soc. London*, pp. 1-13 (1921).
- (44) SPONSLER AND DORE: *Colloid Symposium Monograph* **4**, 174 (1926).
- (45) SPONSLER AND DORE: *Colloid Symposium Monograph* **4**, 194 (1926).
- (46) STAUDINGER, H.: *Ber.* **63B**, 222 (1930) and earlier papers.
- (47) STAUDINGER: *Die hochmolekularen organischen Verbindungen—Kautschuk und Cellulose*. Springer, Berlin (1932).
- (48) WEIMARN, P. P. VON: *Repts. Imp. Ind. Research Inst. Osaka Japan* **7**, No. 5, p. 20 (1926); **9**, No. 5, p. 42 (1928).
- (49) WEITH AND MORY: Chapter in *Kirkpatrick's Twenty-five Years of Chemical Engineering Progress*, p. 213. D. van Nostrand, New York (1933).
- (50) WHITBY, McNALLY, AND GALLAY: *Colloid Symposium Monograph* **6**, 234 (1928).
- (51) WHITBY, McNALLY, AND GALLAY: *Colloid Symposium Monograph* **6**, 232 (1928).
- (52) WILSON: *Colloid Symposium Monograph* **1**, 145 (1923).
- (53) WRIGHT: *Chem. Met. Eng.* **39**, 438 (1932).
- (54) British patents 371,343; 375,424 (1932); 378,074. French patent 742,278 (1933). U. S. patent 1,901,007 (1933).
- (55) British patent 373,406.
- (56) U. S. patent 1,908,546.



# CHANGES IN CRYSTAL STRUCTURE AND THEIR INFLUENCE ON THE REACTIVITY AND CATALYTIC EFFECT OF SOLIDS<sup>1</sup>

J. ARVID HEDVALL

*Chalmers Technical Institute, Göteborg, Sweden*

*Received November 7, 1933*

Without a certain measure of mobility in the structural elements of matter, no measurable reaction between reactive particles can take place. In the gaseous and liquid states this condition is fulfilled at common temperatures, and the mobility of the particles, due to the relative independence of their individual spheres of attraction, forms the very foundation for the peculiarities of these states in several physical and chemical aspects. In most cases it is comparatively simple to define the qualities of particles with mobility of their own in these states of matter. It is possible to use a statistical kinetic treatment, and this, combined with the fact that the experimental work of observing the processes occurring is not especially complicated, makes it natural that the equilibria and velocities of reactions in gases and liquids were investigated first. As for crystals, even their regular external boundaries show that in them we must take into account an inner cohesion caused by fields of force, which of course makes it more difficult to attack them whether by mechanical or by chemical means. An overestimate of these difficulties led to the generalization *corpora non agunt nisi fluida*, which slogan prevented for a remarkably long period interest in the chemistry of the solid state.

A large number of technical processes furnish clear evidence that under favorable conditions solid bodies also possess no small reactivity, provided the other component of the reaction is fluid or gaseous. Such cases are, for instance, the combustion of solid carbon and the reduction of oxide ores; furthermore, such processes as solution or thermal decomposition no doubt indicate reactive solid phases. A typical case of this kind is the combustion of the diamond in oxygen; no one will here imagine either a fused or a gaseous carbon phase as an intermediate reactive body.

At the end of the last century a series of instances were published in which solid phases exclusively were in reaction. I mention the well-known experiments of Roberts-Austen, Violle, Marsden, and several others, who

<sup>1</sup> Translated from the Swedish by S. E. Sterzel, Göteborg.

showed the diffusibility of metals and carbon at relatively low temperatures (23). The contemporaneous discovery of solid solutions and their theoretical treatment by van't Hoff (64) incited increased interest in the chemistry of solid bodies.

At the beginning of this century the rapidly growing work on metallographic problems gave still further impulse in this direction; the pioneer researches of Tammann and his coworkers in particular added many instances of processes in or between entirely solid metal phases.

Omitting the researches of Spring on the reaction between sodium carbonate and barium sulfate, which have been criticized by Tammann (57), the experiments of Cobb (6) on the reaction of calcium carbonate with silicon dioxide may be considered as the first study of a real chemical reaction between non-metallic solid substances; Endell (12) shows foresight when he points out the importance of the study of such reactions for the development of the chemistry of silicate technology.

#### ADDITIVE REACTIONS; EFFECT OF PARTICLE SIZE, ORIGIN OF PREPARATIONS, THEIR HEAT TREATMENT AND CRYSTALLOGRAPHIC CONDITION

When heating cobaltous oxide with a great number of other metallic oxides, it was found that solid solutions of  $\text{CoO-MgO}$ ,  $\text{CoO-MnO}$ , and  $\text{CoO-ZnO}$ , as well as the compounds  $\text{CoO} \cdot \text{Al}_2\text{O}_3$ ,  $2\text{CoO} \cdot \text{SnO}_2$ ,  $\text{CoO} \cdot \text{Cr}_2\text{O}_3$ , and others could form with considerable velocity and completeness even without a flux and at as low a temperature as  $800^\circ\text{C}$ . (19).

These researches were probably the first systematic work on non-metallic additive reactions in the solid state, and showed beyond a doubt that there existed at relatively low temperatures a diffusibility in and between non-metallic solids which at that time had not been expected. The results then obtained by purely chemical means have lately been confirmed through x-ray analyses by Natta and Passerini (53) and by Holgersson (44). During this early work it already was noticed that the previous treatment or origin of the reagents and also their particle size, of course, influenced the intensity or the yield of the reaction. In a series of related researches it was shown both that increased particle size delayed the reaction or necessitated a higher reaction temperature in order to reach a certain yield, and that different primary materials or maximum temperatures in the preparation of the reagents could give products in which various crystal surfaces with different reactivities dominated (22). Modifications of different crystal symmetry are probably very seldom obtained when preparing a substance in different ways (24), but the different atomic spacing in the surfaces has a certain importance. For really different modifications of one substance it was possible to show that those which were richer in energy were the most reactive (21).

This became evident in reactions between silica of different modifications and oxides or carbonates of the alkaline earth metals. The reaction of silica with the oxides is already noticeable after some minutes of heating to about 950°C., and silica can displace carbon dioxide in the carbonates mentioned, below their dissociation temperature at normal pressure. About ten years later Dyckerhoff (10) made a closer examination of the reaction products, viz., the possibility of the formation of different silicates of calcium oxide. The cement and ceramic industries have, as might be expected, also shown a great interest in these and similar reactions. It must be remembered that even though the yield of a reaction between solids has no direct importance, these reactions are nevertheless very important indirectly, because through them the substance changes its composition and therefore also its fusion and sintering temperatures (27).

#### NON-ADDITIVE REACTIONS

The aforesaid reactions were either of the additive type or else, as in the case of silica acting on carbonates, were reactions corresponding to the displacement in fluid phases of a volatile substance by a less volatile one, although the comparison is rather far-fetched with respect to volatility. These latter reactions, however, were the immediate cause of attempts to transfer the "acid anhydride group" in a salt to an added oxide with a stronger affinity for the anhydride, without the anhydride in the interim leaving the solid state. The first experiment in this direction was the heating of barium oxide with calcium carbonate, and the time-temperature curve showed clearly that the reaction had taken the expected turn and resulted in the formation of barium carbonate and calcium oxide. From about 350°C. the curve showed a strong and rather rapid rise of temperature corresponding to a considerable heat of reaction; the horizontal line a little above 900°C. corresponding to the thermal dissociation of calcium carbonate had almost entirely disappeared and was instead found at a higher temperature corresponding to the decomposition of barium carbonate. The reaction was therefore as follows:  $\text{BaO} + \text{CaCO}_3 = \text{BaCO}_3 + \text{CaO}$ . Continued investigations soon showed that this reaction was only an instance of a general but formerly unnoticed type:



or

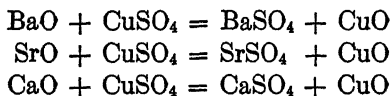


where  $\text{M}^{\text{I}}\text{O}$  and  $\text{M}^{\text{II}}\text{O}$  are two oxides,  $\text{AO}_n$  and  $\text{X}$  are the negative parts of a salt of an oxy acid or other acid, respectively (i.e.,  $\text{CO}_3$ ,  $\text{SO}_4$ ,  $\text{SiO}_3$ ,  $\text{Cl}$ ,  $\text{S}$ , and so on), and  $Q$  is the liberated heat. At first the investigation of



these processes was limited to determinations of those temperatures at which the course of the reaction could be observed on a heating curve, i.e., by observing the temperatures where the reactions set in with greater intensity. Even in the first experiments some peculiarities had been observed which showed that the study of these "reaction temperatures"—which, of course, are not definite temperature constants in the usual sense, but only indicate the more or less marked increase in yield of a reaction which has started at considerably lower temperature—could indeed also lead to an explanation of the mechanism of these formerly unknown reactions. Reliable and consistent results were obtained (32) after developing the method of using heating curves for the study of reactions and of other changes in non-metallic systems. This method had previously been developed for the investigation of thermal dissociations (20) and of exothermic addition reactions involving small evolution of heat. It was now applied to these new reactions, which often were accompanied by a strong evolution of heat (37). Table 1 presents the above-mentioned characteristics of these reactions.

The completeness of these reactions, which happen in a few seconds or less, is attested by the fact that the reactions



gave yields of 87.8, 88.2, and 61.1 per cent, respectively, without regrinding, when an equimolecular reaction mixture was used.

The table shows strikingly that the reaction temperatures are divided into three groups which are well separated. Barium oxide reacts at the lowest and calcium oxide at the highest temperature, as would be expected from their general chemical, thermal, and crystallographic character and as also agrees with the results obtained by Biltz (3). It is also remarkable that the difference between the reaction temperatures of the individual salts of the same kind, e.g., sulfates, is so very small. The influence of the metallic ion is noticeable but not very strong. Nor does the anion of the salt seem to influence the reaction temperature to any great extent *in the salts of the oxy acids used here*, for phosphates, carbonates, silicates, and sulfates react with the same oxide within narrow temperature limits.

It was pointed out in the very first publications that there could be no question of anything but reactions between solids (33). For the reactions with barium oxide, but only for this one oxide, the question might be asked, whether the fusion of the hydroxide or its eutectic could possibly act as an intermediate in the reaction. Balarew (2) put this forward as

TABLE I  
*Reactions of salts with oxides*

SALT	REACTION WITH BARIUM OXIDE		REACTION WITH STRONTIUM OXIDE		REACTION WITH CALCIUM OXIDE	
	Reaction temperature degrees C.	Reaction product	Reaction temperature degrees C.	Reaction product	Reaction temperature degrees C.	Reaction product
Carbonates						
SrCO <sub>3</sub> .....	398 ± 5	BaCO <sub>3</sub> + SrO	—	—	—	—
CaCO <sub>3</sub> .....	344 ± 2	BaCO <sub>3</sub> + CaO	464 ± 4	SrCO <sub>3</sub> + CaO	—	—
MgCO <sub>3</sub> .....	345 ± 5	BaCO <sub>3</sub> + MgO	456 ± 4	SrCO <sub>3</sub> + MgO	523 ± 1	CaCO <sub>3</sub> + MgO
Sulfates						
SrSO <sub>4</sub> .....	372 ± 2	BaSO <sub>4</sub> + SrO	—	—	—	—
CaSO <sub>4</sub> .....	370 ± 6	BaSO <sub>4</sub> + CaO	451 ± 5	SrSO <sub>4</sub> + CaO	—	—
MgSO <sub>4</sub> .....	369 ± 1	BaSO <sub>4</sub> + MgO	441 ± 3	SrSO <sub>4</sub> + MgO	ca. 540	CaSO <sub>4</sub> + MgO
ZnSO <sub>4</sub> .....	341 ± 5	BaSO <sub>4</sub> + ZnO	424 ± 3	SrSO <sub>4</sub> + ZnO	520 ± 6	CaSO <sub>4</sub> + ZnO
CuSO <sub>4</sub> .....	346 ± 0	BaSO <sub>4</sub> + CuO	418 ± 7	SrSO <sub>4</sub> + CuO	516 ± 5	CaSO <sub>4</sub> + CuO
CoSO <sub>4</sub> .....	328 ± 14	BaSO <sub>4</sub> + CoO	431 ± 1	SrSO <sub>4</sub> + CoO	533 ± 4	CaSO <sub>4</sub> + CoO
Phosphates						
Sr <sub>3</sub> (PO <sub>4</sub> ) <sub>2</sub> .....	ca. 350	Ba <sub>3</sub> (PO <sub>4</sub> ) <sub>2</sub> + SrO	—	—	—	—
Ca <sub>3</sub> (PO <sub>4</sub> ) <sub>2</sub> .....	340 ± 5	Ba <sub>3</sub> (PO <sub>4</sub> ) <sub>2</sub> + CaO	ca. 450	Sr <sub>3</sub> (PO <sub>4</sub> ) <sub>2</sub> + CaO	—	—
Pb <sub>3</sub> (PO <sub>4</sub> ) <sub>2</sub> .....	335 ± 3	Ba <sub>3</sub> (PO <sub>4</sub> ) <sub>2</sub> + PbO	453 ± 1	Sr <sub>3</sub> (PO <sub>4</sub> ) <sub>2</sub> + PbO	524 ± 3	Ca <sub>3</sub> (PO <sub>4</sub> ) <sub>2</sub> + PbO
Co <sub>3</sub> (PO <sub>4</sub> ) <sub>2</sub> .....	354 ± 3	Ba <sub>3</sub> (PO <sub>4</sub> ) <sub>2</sub> + CoO	466 ± 2	Sr <sub>3</sub> (PO <sub>4</sub> ) <sub>2</sub> + CoO	ca. 520	Ca <sub>3</sub> (PO <sub>4</sub> ) <sub>2</sub> + CoO
Cr <sub>2</sub> PO <sub>4</sub> .....	342 ± 1	Ba <sub>3</sub> (PO <sub>4</sub> ) <sub>2</sub> + Cr <sub>2</sub> O <sub>3</sub>	464 ± 4	Sr <sub>3</sub> (PO <sub>4</sub> ) <sub>2</sub> + Cr <sub>2</sub> O <sub>3</sub>	517 ± 4	Ca <sub>3</sub> (PO <sub>4</sub> ) <sub>2</sub> + Cr <sub>2</sub> O <sub>3</sub>
Ag <sub>3</sub> PO <sub>4</sub> .....	306 ± 1	Ba <sub>3</sub> (PO <sub>4</sub> ) <sub>2</sub> + Ag + O <sub>2</sub>	461 ± 1	Sr <sub>3</sub> (PO <sub>4</sub> ) <sub>2</sub> + Ag + O <sub>2</sub>	516 ± 4	Ca <sub>3</sub> (PO <sub>4</sub> ) <sub>2</sub> + Ag + O <sub>2</sub>
Silicates						
CaSiO <sub>3</sub> .....	354 ± 1	BaSiO <sub>3</sub> + CaO	454 ± 1	SrSiO <sub>3</sub> + CaO	—	—
(Wollastonite)						
MgSiO <sub>3</sub> .....	354 ± 1	BaSiO <sub>3</sub> + MgO	453 ± 1	SrSiO <sub>3</sub> + MgO	562 ± 1	CaSiO <sub>3</sub> + MgO
(Enstatite)						
MnSiO <sub>3</sub> .....	355 ± 0	BaSiO <sub>3</sub> + MnO	467 ± 2	SrSiO <sub>3</sub> + MnO	563 ± 1	CaSiO <sub>3</sub> + MnO
(Rhodonite)						
Al <sub>2</sub> SiO <sub>5</sub> .....	357 ± 1	BaSiO <sub>3</sub> + Al <sub>2</sub> O <sub>3</sub>	429 ± 2	SrSiO <sub>3</sub> + Al <sub>2</sub> O <sub>3</sub>	532 ± 1	CaSiO <sub>3</sub> + Al <sub>2</sub> O <sub>3</sub>
(Sillimanite)						

his opinion, based on some experiments in which the necessary care for obtaining reliable temperature measurements had in no way been exercised. Without mentioning here all the evidence for the defects in Balarew's working arrangements and the consequent errors in his conclusions, which were immediately pointed out by several authors (36, 47, 59), it may be mentioned here that barium hydroxide reacts less intensively and often at higher temperatures than does barium oxide. This might be expected because the attraction of the oxide to the acid constituent of the salt indubitably depends on the same affinity for it as for acid anhydrides in the gaseous state, e.g., carbon dioxide; and this attraction ought to have a certain connection with the degree of unsaturation of the attracted substance.

Nor, as was also shown, could a gaseous anhydride, such as carbon dioxide or sulfur trioxide, be the cause of the reaction. A comparison between the reaction temperatures and the vapor tension showed the impossibility of this. Indeed in several cases it was possible to determine the relatively unimportant rôle that was played by the gaseous phase at the higher temperatures which were a consequence of the exothermic reactions in the solid state (41). In such systems as mixtures containing silicates or some phosphates (cf. table 1), it is of course just as nonsensical to talk about a reactive gaseous phase as it is to try to find a liquid phase in the reaction mixtures containing strontium oxide or calcium oxide, where strontium hydroxide or calcium hydroxide decompose without previous fusion. It may be added that only on the assumption that these reactions occur in the solid state can they be explained from a common point of view. This ought *a priori* to have given less appeal to the search for other explanations, conceivable in some special case, perhaps, but not generally possible. Continued research in this field has not been promoted by these opinions in any way other than by compelling us to increase the accuracy of our experiments still more. They have been mentioned in their historical relation, because at the very start they were questions that craved an answer.

Of course, a fusion when caused by impurities or by the reaching of a fusion point or a eutectic temperature through the heat of reaction may in some cases increase the yield of the reaction. According to what has been said before, this can also occur through the liberation of a gas during the heating. This increase in yield which can be caused by direct reaction with a liquid or gaseous phase can also be caused by these phases in destroying or decreasing the thickness of the diffusion layer of the reaction product, thus producing new contact surfaces. At first considerable attention was given to these phenomena (35). Recently we have published the results of experiments intended to determine what quantities of a fusion caused by

added impurities were necessary to increase the yield of the reaction to any appreciable extent. The quantities so found considerably exceeded the amount of impurities which can be expected in preparations of high quality (28). It was also found that it is not necessary that the appearance of fusion should always help a certain reaction (26). In reaction mixtures such that low fusing compounds or eutectics are formed on account of the original composition, or through the reaction, it is always easy to determine from which temperature the influence of a liquid phase has to be taken into consideration; for even in cases where the eutectic or polyeutectic conditions are not known it is always possible to find when fusion starts by determining the electrical conductivity (26, 60).

No doubt the results from a new field of science should be met with healthy doubt, but it is interesting to see how in certain quarters there was an extreme unwillingness to believe in the possibility of reactions between solid substances only. Although not mentioned in the chemical literature by anybody except by Balarew it was commonly thought that something like a catalytic action of extremely thin liquid surface films was required, and even that those small traces of gas which correspond to the extremely small vapor tensions of substances at temperatures several hundred degrees below their dissociation point, could cause this reactivity. Everybody who has worked with these reactions and has studied their high velocity must consider the possibility of a gaseous phase as the reaction medium as practically impossible. Certainly this is so at the beginning and in the first stages of the reaction, even when only sulfates or carbonates are considered and when no consideration is given to the absolutely impossible case of a gaseous phase in the reaction with the silicates. Tammann, who together with his coworkers gave a number of instances of this type some years ago, has found one such reaction which is particularly favorable, viz.,  $\text{PbO} + \text{CuSO}_4 = \text{PbSO}_4 + \text{CuO}$  at about  $500^\circ\text{C}.$ , where the sulfate has a vapor tension of 37 mm. Even here the gas phase had very little influence on the reaction yield (59).

It seems that these attempts to make a gaseous phase act as a medium for the reaction were based on a conception which was borrowed from the well-known hypothesis put forward by Nernst after Debray's researches on the thermal equilibria of solids, and which was intended to make possible the application of the Guldberg-Waage law to heterogeneous equilibria. But it has been overlooked that this conception is entirely incapable of explaining rapid reactions. Against the assumption of the existence of thin active liquid layers it is of course not sufficient to show that absolutely dry materials have been used. It must rather be shown *that the courses of the reactions are connected with and vary with factors which evidently are characteristic of the solid state.* Such a working plan was, indeed, used

from the very beginning of this investigation. It had to be, because any other attempt to use *one* explanation for these reactions was found to be impossible. The results obtained have also fully proved the truth of the hypothesis that *every interference with the crystal lattice which sufficiently decreases its interior stability*, increases the mobility of the units of the lattice, and therefore more or less stimulates the crystal to react with another crystal.

#### THE INFLUENCE OF WEAKENING THE COHESION OF THE LATTICE

##### 1. *Heating and deformation*

Factors which widen or weaken the crystal lattice are: above all, heating; changes in the symmetry of the lattice; and the introduction of foreign substances into it. Already in our first publication (31) it was assumed that the ionic lattice which salts of some oxy acids contain at low temperatures could partly change when heated to an anhydride lattice, hence in a sulfate to metallic oxide and sulfur trioxide, the extent of change being dependent on the temperature. This should be an intermediate state between the low temperature ionic lattice and thermal dissociation at high temperature. Such a change must be regarded as extremely probable, considering what precedes the thermal decomposition of the lattice, and has lately been verified by measurements of the conductivity and diffusion made by W. Jander (49). X-ray photographs at different temperatures, which should give interesting results in this case, do not, as far as I know, exist.

In subsequent publications on non-additive reactions, the fact that barium oxide reacts at the lowest temperature, being followed by strontium oxide, calcium oxide, and so on (25), was explained on the basis of the above-mentioned working hypothesis. It need only be mentioned here that it seems quite natural that barium oxide, having a stronger affinity to acid anhydrides, ought to cause a reaction at a lower temperature and a correspondingly lesser deformation of the ion group (e.g.  $\text{SO}_4$ ) or a smaller number of free or loose anhydride complexes per unit volume than do strontium oxide or calcium oxide. It was also shown that the very small influence of the metallic ion of the salt of an oxy acid on the reaction temperature was in good agreement with the great independence of anions of this kind revealed by Schäfer and Schubert (56) through measuring their infra-red absorption. Further confirmation of this conception was given by Fajans (13), working on the anions of oxy acids, who showed that these anions are very non-susceptible to the deforming action of cations. Taking into account the dominating importance of the atoms of oxygen in building up the crystal lattices of the silicates and oxides, as

pointed out by W. L. Bragg, it seems possible that this is also responsible for the independence of the anions in the salts of oxy acids (5). Such a connection might possibly be able to explain the similar reactivity of salts of different oxy acids.

On further study of reactions of this or similar types, diagrams were made showing the reaction yields at different temperatures but with constant time of reaction. It was then found, as expected, that a slow reaction started at considerably lower temperature than the rapid progress indicated by the heating curves. Curves were obtained of the same type as those for other properties which depend on the inner mobility of the lattice, such as conductivity, vapor tension, etc. With properly selected reaction periods these curves showed discontinuities at the same temperatures as do the corresponding heating curves. It was consequently evident that the thermal expansion of the lattice, by increasing kinetic energy of the particles and their consequent momentary opportunities for moving outside their mutual fields of attraction, had caused the expected increase in reactivity.

If the crystal lattice of the salt in question contains ions which can exert deforming or polarizing action, it can be expected that the above-mentioned purely thermal expansion at lower or higher temperatures ought to be changed by deformation effects. In an ionic lattice a partial de-ionizing at various points in the lattice could thus occur to a certain extent. Such changes of course imply disturbances in the homogeneity of the field of force in the ideal lattice, which cause a local weakening and facilitate mobility of the particles or transport of matter. These consequences of the action of the ions on one another should be observed most readily in ionic lattices with strongly deforming cations (small and of non-rare gas type) and easily deformable anions (as already stated, not anions of oxy acids but halide ions, and especially the larger ones).

One must consider the incomplete extent to which it is possible to survey the phenomena of deformation, the thermal data of the crystal lattices (lattice energy, heat of formation) and the degree of ionization as functions of the temperature, and must also keep in mind the probable difference in action between ions and uncharged particles as reaction partners, not only for different crystal combinations but also for the same mixture at different temperatures. Hence it cannot be expected that heating curves alone should give any *quantitative* information even of a relative nature about the inner condition of the lattices or about the character of the reacting particles. Even the simple and approximate heating-curve method, however, gives *qualitative* results which are evidence for the correctness of the above-mentioned opinions. Thus the sulfates whose  $\text{SO}_4$  groups, according to Fajans, are very resistant to the action of deforming cations,

form a reaction system much more homogeneous in reaction temperatures than the halides, the ions of which are much more sensitive to deforming influences (table 2).

In the case of the sulfates the temperatures are grouped, as would be expected from lattice energy data and from the possibility of deformation of the sulfate ion, which, as aforesaid, is very small and therefore *a priori* makes a stronger influence of the cation of the salt improbable. On the other hand, within the halide group there is strong individuality in the reacting salt, caused by a closer influence of the ions on one another.

TABLE 2  
*Reaction temperatures of sulfates and halides*

SALT	REACTION TEMPERATURE WITH		
	BaO	SrO	CaO
	<i>degrees C.</i>	<i>degrees C.</i>	<i>degrees C.</i>
SrSO <sub>4</sub> .....	372	—	—
CaSO <sub>4</sub> .....	370	451	—
MgSO <sub>4</sub> .....	369	441	540
ZnSO <sub>4</sub> .....	341	424	520
CoSO <sub>4</sub> .....	328	431	533
CuSO <sub>4</sub> .....	346	418	516
Ag <sub>2</sub> SO <sub>4</sub> .....	342	At transition point of Ag <sub>2</sub> SO <sub>4</sub> (Cf. table 4)	
CuCl.....	270		
CuBr.....	312		
CuI.....	340		
PbCl <sub>2</sub> .....	273		
PbBr <sub>2</sub> .....	248		
PbI <sub>2</sub> .....	ca. 200		
NiCl <sub>2</sub> .....	312		
NiBr <sub>2</sub> .....	272		

## 2. Foreign constituents in the crystal lattice

Very similar to this kind of weakening of the cohesion of the lattice is the exchange of a small part of one or more of the units of the crystal lattice, as is the case in isomorphous mixed crystals or in solid solutions. Differences in the radii of atoms or ions, or differences in electrical charge between the original and the foreign particles of the lattice can, when the proportion between them is favorable, cause internal changes which should produce a change in reactivity compared with that of the fully uniform crystal. Several investigations of such cases have proved the correctness of this hypothesis (26). Table 3 shows this for reactions between barium oxide and in one case pure lead chloride; in another case a mixed crystal of lead

chloride + 0.1 molar per cent of barium chloride. The foreign substance in this case being the same as the compound formed in the reaction, no complicated eutectics or reactions are obtained, and the possibility of a fusion is entirely eliminated at the temperature in question. According to Goldschmidt (18), the ionic radius of  $\text{Ba}^{++}$  is  $1.43 \times 10^{-8}$  cm., and of  $\text{Pb}^{++}$   $1.32 \times 10^{-8}$  cm.

Continued investigations of the same kind with this and several other reaction mixtures have shown that these additions also influence such characteristics as surface formation, particle size, velocity of solution, and catalytic effect. It has also been ascertained that there exist limits for the admixture of foreign substances beyond which a contrary effect obtains, and that this is often the case at rather low percentages. These results are

TABLE 3  
*Reactions in pure and impure crystals*

REACTION TEMPERATURE	PER CENT OF $\text{PbCl}_2$ REACTED ACCORDING TO EQUATION $\text{BaO} + \text{PbCl}_2 = \text{BaCl}_2 + \text{PbO}$	
	With pure $\text{PbCl}_2$	With $\text{PbCl}_2$ + 0.1 molar per cent of $\text{BaCl}_2$
<i>degrees C.</i>		
228.0	10.6	15.5
235.0	11.6	22.0
248.0	16.8	24.0
258.0	23.9	27.0
269.0	27.5	30.1
279.0	29.2	33.5
289.0	34.3	43.3
292.0	39.8	49.5
298.0	45.0	55.4

of interest in connection with the production and use of mixed catalyzers and their activity. Investigations by the writer and his coworkers have shown that even very small admixtures (0.1 molar per cent) of manganous oxide, magnesium oxide, or nickelous oxide in cobaltous oxide appreciably decrease the size of the cobaltous oxide crystals and increase their tendency to oxidize to  $\text{Co}_3\text{O}_4$ . Similar results with regard to particle size have also been obtained by Wyckoff (66) and by Edner and Schönfeld (11).

### 3. *Different kinds of structural changes*

Nearly ten years ago the first instance of a method of greatly increasing the reactivity of the solid state was published, a method which, because of its manifold applications and practical importance, seems to be of general interest (34). Starting from the same working hypothesis as before, the



supposition was made that *a crystal which at a given temperature sustains a change of phase in the form of a polymorphous transition or in any other form, ought to show, during this process when the particles have an abnormally great freedom of movement, an abnormally great reactivity when mixed with another reactive substance, whether solid, liquid, or gaseous.* Continued and still pending researches in this field have shown that the reactivity is a very sensitive indicator of every kind of such change of phase. A great number of crystallographic changes, both between different systems and classes and in the so-called secondary structure, including the fusion of a substance, magnetic transitions without change of lattice structure, and changes of phase of alloys (e.g., when a solid solution at a given temperature separates into two), all provide instances of such an effect. In what follows some of the results obtained in this field will be mentioned.

The characteristic fact has already been mentioned that when basic oxides react with salts of oxy acids the oxides react with a group of salts

TABLE 4  
*Reactivity and transition temperatures*

OXIDE	WITH SILVER NITRATE		WITH SILVER SULFATE	
	Reaction temperature	Transition point	Reaction temperature	Transition point
	<i>degrees C.</i>	<i>degrees C.</i>	<i>degrees C.</i>	<i>degrees C.</i>
BaO.....	170 $\pm$ 1	160	342 $\pm$ 4	411
SrO.....	172 $\pm$ 3		422 $\pm$ 1	
CaO.....	164 $\pm$ 5		422 $\pm$ 5	

within a temperature range which is characteristic for the oxide, so that barium oxide reacts at about 350°C., strontium oxide about 100°C. higher, and calcium oxide at about 70°C. still higher. Yet the reactions in mixtures of each of these oxides with silver nitrate or silver sulfate, which have transition points at 160°C. and 411°C., respectively, showed the reaction temperatures listed in table 4.

For silver nitrate the transition point from rhombic to rhombohedral lies at 160°C. The reaction temperatures of all three oxides lie immediately above this temperature, when the change has reached the necessary velocity. The influence of the corresponding change in silver sulfate is entirely analogous, although, of course, it cannot affect the reaction temperature with barium oxide, which normally lies lower than the transition point. Similar behavior has been proved for the reaction between barium oxide and silver iodide, according to the equation  $\text{BaO} + 2\text{AgI} = \text{BaI}_2 + 2\text{Ag} + \frac{1}{2}\text{O}_2$ . In this case heating curves as well as yield curves have been

obtained. Figure 1 shows that the yield reaches a maximum where the change from hexagonal to regular shape is rapid, and therefore decreases with increasing temperatures (39).

To this class of changes belong also those changes of phase of silicon dioxide which have a considerable influence on its reactivity. According to Fenner (14) there occurs at  $575^{\circ}\text{C}$ . a less radical displacement of the atoms accompanied by a change of symmetry, in which silicon dioxide is rapidly changed from  $\beta$ -quartz to  $\alpha$ -quartz or the reverse. On continued heating nothing happens until  $870^{\circ}\text{C}$ . when the  $\alpha$ -quartz becomes unstable, after which, in the absence of a flux, no tridymite but instead cristobalite is formed. On heating quartz with  $\text{Fe}_2\text{O}_3$  in oxygen (to prevent reduction), the yield curve shown in figure 2 was obtained for heating periods of 2 hours each. The reaction product was rose-red and contained up to 4.5

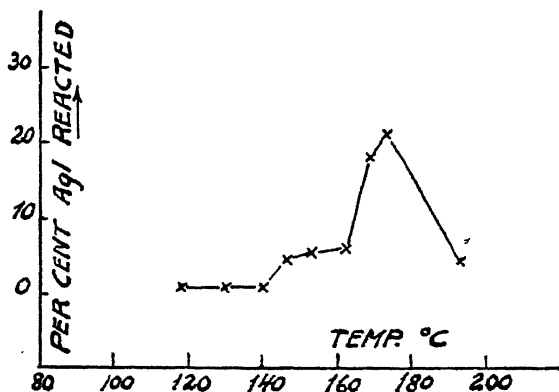


FIG. 1. REACTION VELOCITY AND TRANSITION TEMPERATURE

per cent of ferric oxide which was insoluble in acids. From the curve, it is evident that this reaction starts, though only to a slight extent, at the first rapid transition, then stops until about  $900^{\circ}\text{C}$ . when the  $\alpha$ -quartz changes slowly to cristobalite, giving the ferric oxide plenty of time to react. The x-ray photographs showed that the penetration of ferric oxide into the silicon dioxide lattice gave a displacement of the cristobalite lines corresponding to an expansion of the lattice (43).

It is evident that this quite unexpected solid solution of ferric oxide *can only be formed in connection with transition processes*, for neither cristobalite nor tridymite, when mixed with ferric oxide, gives the red product, nor do  $\beta$ -quartz or  $\alpha$ -quartz do so below their transition points (43). Besides, the rose-red product formed is of interest in itself, because this state of the iron oxide, which seems rather unnatural from a crystallo-

graphic point of view, causes widely changed characteristics both of ferric oxide (if it can be said to exist individually in the solid solution) and of the product compared with pure cristobalite, especially with respect to its catalytic activity.

From an investigation with S. Ljungkvist (30b), the following facts can be briefly stated. After a method had been devised to obtain homogeneous products containing no free ferric oxide, two solid solutions of ferric oxide in cristobalite were made, one containing 0.24 and the other 1.0 per cent of  $\text{Fe}_2\text{O}_3$ . These preparations were tested with regard to the sensitiveness of the color to change of temperature, the possibility of reducing the dissolved ferric oxide, and their catalytic activity on ethanol vapor. Whereas pure ferric oxide made from the oxalate began to darken

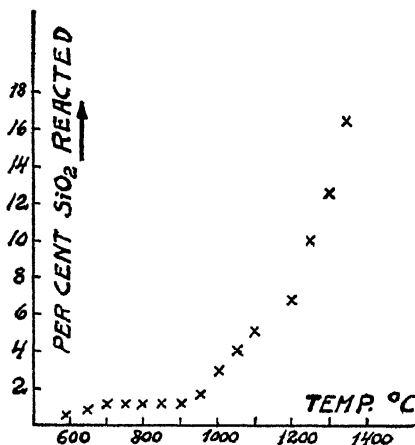


FIG. 2. REACTIVITY OF SILICA

at about  $300^\circ\text{C}$ ., such a change already sets in with the red  $\text{Fe}_2\text{O}_3$ - $\text{SiO}_2$  products at  $240^\circ\text{C}$ . The change begins quite suddenly and reaches a grey-to brown-violet final stage at  $270^\circ\text{C}$ . in contrast with the behavior of pure ferric oxide, which darkens to a deeper and deeper color at far higher temperatures. These color changes are reversible in the red products. Whereas in respect to color changes the dissolved ferric oxide must consequently be considered much more sensitive to heating than is pure ferric oxide, the opposite is the case, as could be expected, as regards resistance to chemical reagents. It is neither dissolved by acids nor by solutions which could form complex ions containing iron, and it is also much more resistant to reducing gases. Hydrogen reduces free iron oxide at  $300^\circ\text{C}$ . but leaves the red solid solution unattacked even after heating for 1 hr.

at 400°C. Saito (54) has found that ferric oxide is reduced by a gas mixture containing 5 per cent of carbon monoxide at and above 300°C.; the red products were entirely intact, however, after heating for three-quarters of an hour in a mixture of 25 per cent carbon monoxide and 75 per cent carbon dioxide at 580°C. From a catalytic point of view the considerably increased stability in the presence of vapors of ethanol or its decomposition products are also of great interest. Pure ferric oxide is reduced below 300°C.; the solid solutions are on the contrary unchanged even at 440°C. Experiments with R. Hedin (30b) have also shown that the solid solutions have a greater reactivity in the solid state, e.g., when reacting with calcium carbonate, than does pure cristobalite (cf. figure 3).

The catalytic activity, at least in such reactions as have been investigated, is greatly changed when compared with that of ferric oxide or of cristobalite when all have had the same treatment. Thus, for instance, the breaking down of ethanol with the formation of aldehyde, ethylene, ethyl ether, and hydrogen, already occurs at about 400°C., whereas cristobalite does not act until about 60°C. higher. On the other hand free ferric oxide shows quite different behavior. As before mentioned it is reduced at a much lower temperature. For selective catalysis, especially with regard to the formation of heat-sensitive products or the stability of the catalyst, these characteristics should command interest (30b).

A similar maximum of reactivity at the transition point has been obtained for the change of rhombic to monosymmetric sulfur in an investigation with A. Floberg and P. Paulson, the results of which are in press (28a).

To reaction effects caused by structural changes can be assigned the increased intensity of reactivity of bismuth and copper at 75°C. and 70°C., respectively, as shown by a recently completed investigation by the writer with R. Hedin and S. Anderson. In the literature on measurements of the physical constants of these metals, especially of bismuth, one observes that opinions are very much divided on these "transition points," the existence of which has been found by Cohen (7, 8) and his coworkers. A tabulation of the results obtained with different measuring methods and by different investigators leads to the opinion, that here we are dealing with phenomena which certainly exist at the given temperatures, but can hardly be characteristics of the primary crystal lattice. That this is the case, we have also shown by Laue diagrams above and below the "transition point," which do not show any change of structure. The appearance or non-appearance of these "transitions" seems indeed to depend to a remarkable degree upon accidental circumstances during the formation of the crystal. As it was of interest to find out if the rule of increased reactivity during transitions could also be applied to changes of this kind, determina-

tions were made of the velocity of solution of bismuth and copper in dilute nitric acid and for bismuth also in an iodine-potassium iodide solution (figure 5). The tests were made in a thermostat at constant temperature ( $\pm 0.05^\circ\text{C}$ ). The reaction time was 30 minutes for bismuth and 60 min-

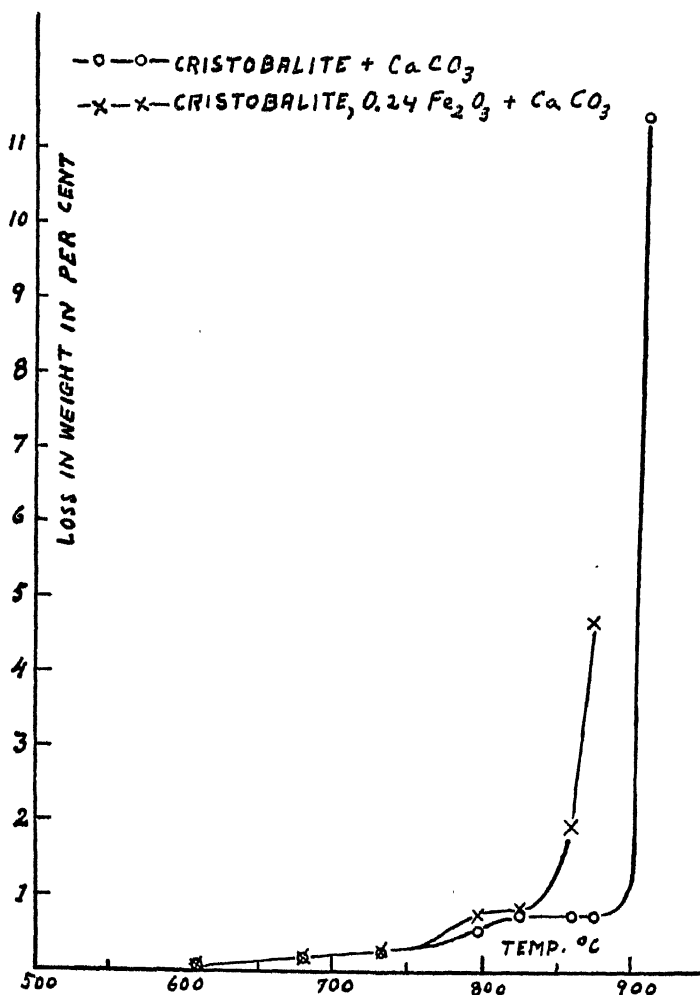


FIG. 3. REACTIVITY IN THE PRESENCE OF SOLID SOLUTION

utes for copper, and was determined after dilatometric tests had given the time of change in the test pieces used. Results were obtained which were entirely in agreement with the temperatures reported by Cohen. The transition period for bismuth was 30 minutes at rising temperature, and

90 minutes at falling temperature; for copper in both cases it was about 60 minutes. Only a brief résumé of the results of this investigation will be given here, as a fuller report will shortly be published. As can be seen from figure 4, a considerable increase of intensity in the reaction with nitric acid is found in Cohen's "transition range." In the series of tests with rising temperatures the increase begins at about 74.1°C. and reaches a maximum at 74.5°C., corresponding to the maximum between 74.9 and 73.5°C. on the curve with falling temperatures. The differences between the two curves at temperatures above the "transition point" depend upon the fact that the test pieces for the curve with rising temperatures, when heated

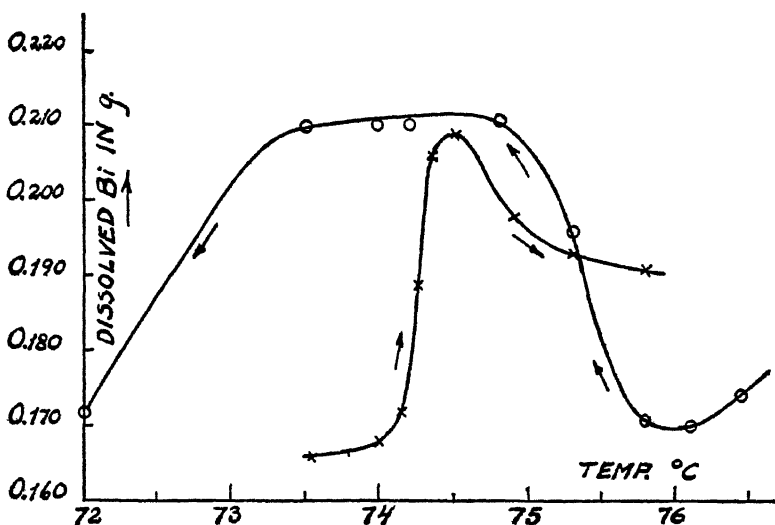


FIG. 4. RATE OF SOLUTION OF BISMUTH IN NITRIC ACID AT TRANSITION TEMPERATURE

in the acid to the test temperature, rapidly pass the "transition range," which consequently has a certain influence in this case also. The circumstances are analogous when copper dissolves in nitric acid or when iodine reacts with bismuth.

In a recent publication Goetz and Hergenrother (17) have proved that in this case it cannot be a question of transitions in the usual sense, but instead these are changes in a kind of secondary structure which does not belong to the lattice, but to the macrocrystal, and which is caused by a certain periodicity (depending on the chemical character of the crystal and its temperature) in the formation of contracting and more resistant atomic planes, which divide the crystal into a block- or mosaic system. Such phenomena, which certainly have a close connection with the forma-

tion of etching figures, have been the subject of a number of recent investigations, especially by Zwicky (68) and Goetz (16).

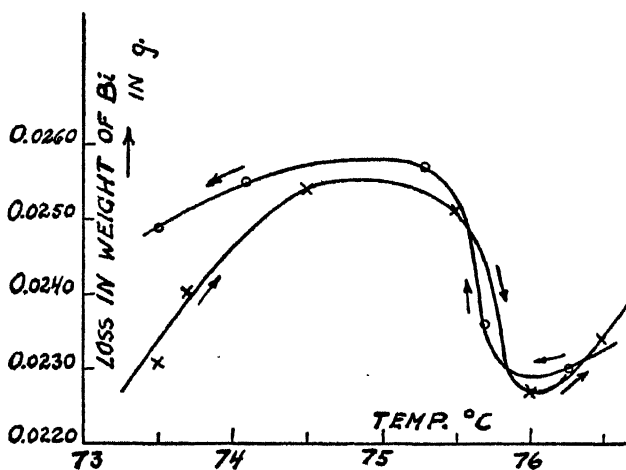


FIG. 5. RATE OF REACTION OF BISMUTH WITH IODINE AT TRANSITION TEMPERATURE

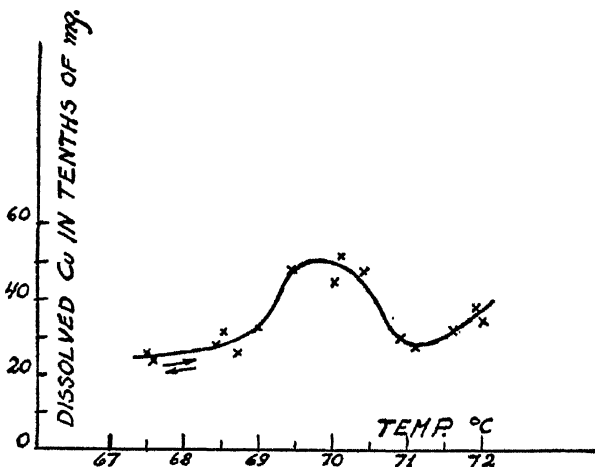


FIG. 6. RATE OF SOLUTION OF COPPER IN NITRIC ACID AT TRANSITION TEMPERATURE

#### 4. Changes in the crystal lattice through radiation

Every absorption of energy quanta which increases the energy of an atom, an ion, or molecule, or of the crystal lattices built up of these, ought to lead to a change of reactivity in case of sufficient radiation and sufficient quantum yield. Especially in gaseous or liquid systems several instances

of such photochemical effects produced by visible or ultra-violet radiation have long been known (cf. Dhar, *The Chemical Action of Light*). A number of such phenomena are also known in crystals. In the first place, of course, in the silver halides used in photography, a de-ionization with formation of uncharged Ag and Br particles occurs in certain places such as in the AgBr ionic lattice. There can hardly be any doubt that such a change in the lattice of the silver bromide, being to some extent comparable to an admixture of foreign particles, ought also to change the reactivity between silver bromide and another suitably selected reactive crystal, e.g., a solid oxide as described above. *Moreover, such an effect can also be expected in those substances in which the change on irradiation does not reach the stage of ionization or de-ionization, but leads only to electron displacements which produce phenomena of phosphorescence or possibly also of fluorescence.* As far as I know, investigations of this kind have not heretofore been made, and the work in progress here has as yet not given results ready for publication. We have started, however, to make comparisons of reactivity and catalytic effect between such differently sensitized or sensitive substances in normal and in activated condition, using irradiation with visible or invisible radiation, including x-rays.

Some difference between visible and x-ray radiation seems to be probable, as it is known from the investigations of Glocker (15) that the chemical action of x-rays is connected with the formation of more or less rapid photoelectrons or in certain cases Compton-electrons, and therefore should be more of a secondary nature than the usual photochemical processes. The Einstein photochemical equivalence law cannot reasonably be expected to be valid, but there should always be some proportion between "real absorption" (photoelectron formation) and chemical action. Since a photoelectron, given a higher or lower velocity, can attack, and eventually in a certain reaction can activate or inactivate a number of particles in the lattice, thus disturbing the uniformity of the interior cohesion of the lattice, it is evident that to attain a perceptible result it is possible to take into account the change not only of that particle from which the electron has been thrown out, but of a considerably greater number. (It has been calculated that one thousand silver bromide "molecules" can be attacked for each absorbed quantum of x-ray radiation.) For the demonstration of this possibility, it should be advantageous to use substances with a strong x-ray absorption; for, contrary to the case with light, a relatively small absorption, i.e., a low efficiency, has to be counted on. Such substances do not appear to have been used.

In exothermic reactions it is also desirable to determine whether a quite small addition of energy through radiation may not be able to cause chain reactions,—at least so-called energy chains. The rigidity of the solid



system and the difficult or delayed contact between the reaction components in the solid mixture are perhaps in most cases obstacles serious enough to prevent the establishing of such an effect, even when using sources of radiation giving more energy than those which were at our disposal in the previously mentioned experiments. *A priori*, experiments with exothermic rearrangement or decomposition inside the crystal lattice of a single substance look more promising. Experiments of this kind, e.g., increasing the velocity of transition of an undercooled high-temperature modification or the decomposition of an endothermic compound, deserve closer investigation.

*Endothermic reactions*, as also the thermal decomposition of exothermic compounds, should, however, be capable of being affected by suitable irradiation. If the dissociation energy is supplied by heat, it should be possible to change the velocity of the decomposition. In such cases, of course, there can be no question of chain reactions, but the action of the photoelectrons on the normal electric uniformity of the lattice or on the momentary energy state of certain particles, should be able to influence the internal diffusion in the crystal, if the substance and radiation are suitably selected, and therefore affect also the velocity of thermal decomposition. Such preliminary experiments have been made with compounds of lead and barium as shown below.

For the investigations made by the writer in collaboration with L. Nelson salts of heavy metals have been selected. The objection may be raised that the velocity of the photoelectrons may decrease with increasing nuclear charge, and thus the gain through strong absorption may be compensated or overcompensated. Yet, according to Bothe (4) and to a private communication from Glocker, it does not seem permissible to postulate a simple connection between the nuclear charge and the velocity of photoelectrons of a certain radiation, nor has it been possible with x-rays to find such a selective absorption of certain frequencies as is the case with chemically active light. In preliminary experiments—and practically all experiments on the chemical action of x-rays are at present preliminary, for in this field the conceptions are still more muddled than in the rest of photochemistry—it is therefore not even possible to recommend the use of very homogeneous radiation.

In our experiments we used a Müller-metalix tube ("media") giving tungsten radiation. The voltage used was 60 to 70 kv. and the current 2.5 to 3.0 ma. The tungsten anticathode was about 17 cm. from the irradiated powder, which was kept in a small quartz flask open in the one constricted end to permit the escape of gas and the introduction of the wires of a thermocouple. During an experiment two such flasks with the same substance, one of them protected from the radiation, the other one

exposed to it, were rotated at uniform speed in the heating apparatus standing immediately below the x-ray tube. This apparatus consisted of an electrically heated furnace with a long constant-temperature space and of large heat capacity, which could be kept at a constant temperature ( $\pm 1^\circ\text{C}.$ ) within the space where the flasks were placed. As the time required to reach the reaction temperature is of great importance in its influence on the course of the reaction, and as this could not be the same for the irradiated and the protected substance on account of the change in heat capacity which is caused by the material designed to absorb the radiation, the rotating device was placed in the furnace inside a stationary brass tube with a wall thickness of 5 mm. This tube had an opening at the top directly over the flask to be irradiated, and to equalize the heat capacity an opening of the same size on the lower side, under the protected flask. As the radiation penetrated the brass tube to a certain extent, the non-irradiated flask was further protected by a lead pipe 2 mm. thick, or for high temperatures, by a steel pipe 5 mm. thick outside the brass tube, with an opening corresponding to that of the brass tube. Using this arrangement it was possible to obtain a synchronous and absolutely constant rate of heating during any period for both irradiated and non-irradiated samples. A series of experiments with parallel concentrated ultra-violet radiation from a mercury-arc lamp of Kromayer type was also made.

The reaction time in these experiments was varied within wide limits, because in the case of chain reactions of the above-mentioned type, positive results could possibly be expected even with rather short exposures. The radiation from the x-ray tube had to pass about 3.5 mm. of quartz in all before reaching the powder.

The results were negative throughout, possibly because of the use of too weak a source of energy. In the exothermic reaction,  $\text{CaO} + \text{NiCl}_2 = \text{CaCl}_2 + \text{NiO}$ , which had been the subject of an earlier research of the writer and Nelson (40), there was no distinct difference in the yield between the irradiated and the non-irradiated sample when ultra-violet radiation was used for 15 minutes at  $130^\circ\text{C}.$ , nor with x-radiation at 61 kv. and 2.5 ma. for 15 minutes at  $139^\circ\text{C}.$ , nor with the same x-radiation for 180 minutes at  $18^\circ\text{C}.$

In the endothermic reaction,  $\text{PbCO}_3 = \text{PbO} + \text{CO}_2$ , x-radiation of 70 kv. and 3.0 ma. was used at  $194^\circ\text{C}.$ ,  $206^\circ\text{C}.$ , and  $269^\circ\text{C}.$  without difference in yield under the radiation. The lead carbonate was made by the method of Centnerszwer and Awerbuch (9).

In another endothermic reaction,  $4\text{PbO}_2 = 2\text{Pb}_2\text{O}_3 + \text{O}_2$ , studied both under x-rays and under ultra-violet radiation and at temperatures between  $306^\circ\text{C}.$  and  $315^\circ\text{C}.$ , with the time varied between 15 and 270 minutes, the

differences were somewhat larger, but so also were the experimental errors, so that no definite influence of the radiation was ascertainable.

Finally the reaction,  $\text{BaO}_2 = \text{BaO} + \frac{1}{2}\text{O}_2$  was selected because of the strong absorption of x-rays by barium and because of the binding between the atoms of the peroxide. X-radiation of 65 kv. and 3.1 ma. was applied at 575°C. for periods up to 300 minutes but without evidence of influence by the radiation. Mixtures of barium peroxide and aluminum oxide also failed to show any such influence.

These investigations will be continued with better equipment in order to find out what really happens when x-rays act chemically. This can best be done by the study of the decomposition of solids, using simple kinds of atoms, ions, or lattices, or characteristic types of compounds. It seems unnecessarily involved to undertake this important study with solutions of complicated compounds, with complex organic substances, or by color changes in minerals, as has been done heretofore.

#### TRANSITION POINTS AND CATALYTIC ACTIVITY

The importance of widening the crystal lattice or of rendering it inhomogeneous, which is well-known in the use of "mixed catalysts," has been mentioned in a previous section. In certain cases, however, there should be still one more possibility, as yet not described, namely to change the activity of a catalyst by allowing it to act within the temperature range corresponding to a rapid reversible change in it. Experiments of this kind are now being made with repeated passage through the transition range in both directions. The small fluctuations of temperature can be obtained by surrounding the catalyst chamber with a heating element, the heat capacity of which is small compared with that of the furnace, and which is periodically shunted into the circuit.

It should also be mentioned that experiments by the author and E. Gustafsson indicate that magnetic transitions give similar effects (30a). Using nickel to catalyze the formation of carbon dioxide from carbon monoxide, or the decomposition of nitrous oxide, discontinuities in the yield are obtained at the magnetic transition point (358–360°C.), as shown by curve 1 in figure 7. The sudden increase in gas volume at the transition point can only be explained by a more rapid decomposition following the equation  $\text{N}_2\text{O} = \text{N}_2 + \frac{1}{2}\text{O}_2$ . Curve 2, where pure nitrogen is passed through the tube, shows that no discontinuity can be observed due to increase in volume of nickel at the transition point. The gas volume introduced into the tube was the same in both cases. Such an effect is, however, highly selective, as could be expected. It should be mentioned that the extended application of the law of increased reactivity during structural changes has recently been applied to catalytic reactions by Hüttig (45), who thus was able to

explain that change in catalytic activity which can be observed during the so-called aging period of the catalyst.

REACTIONS IN ORE ROASTING AND THE INFLUENCE OF THE NASCENT  
STATE IN SOLIDS

In connection with certain studies on roasting and other metallurgical processes, some investigations were made some years ago on those rapid reactions which take place when oxides of alkaline earths are heated with

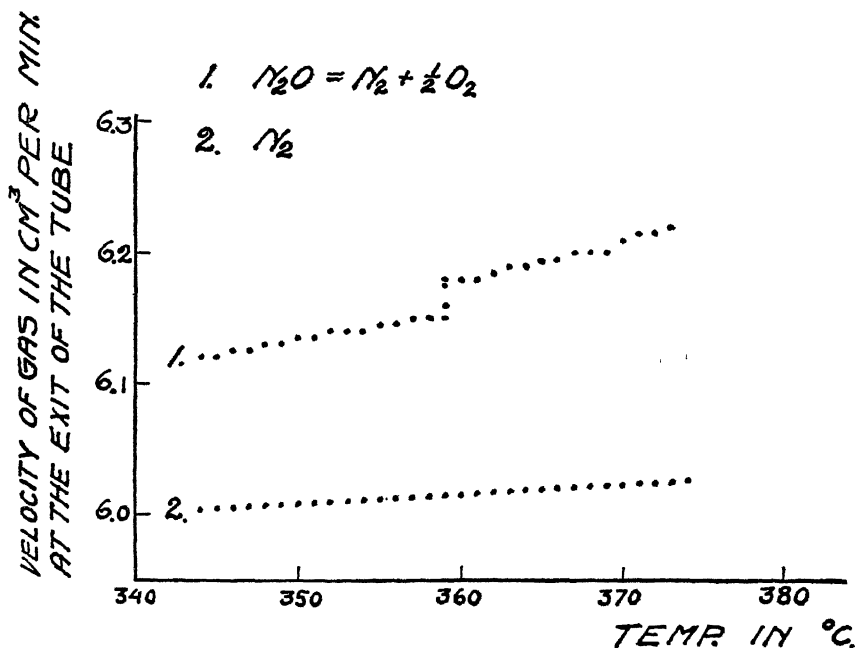
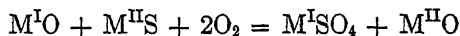


FIG. 7. CATALYTIC ACTIVITY AT THE MAGNETIC TRANSITION POINT

metal sulfides in air or oxygen. These reactions follow the general equation



and it was shown analytically that the rôle played by gaseous sulfur dioxide or sulfur trioxide was very small compared with the direct reaction between solid oxide, solid sulfide, and oxygen gas. Continued investigations showed also that these reactions are instances of a general principle; for just as sulfides reacted to form sulfates, so phosphides give phosphates,

carbides give carbonates, and silicides give silicates of the metal of the added oxide (42). In metallurgical practice such reactions have often been used without their nature being known. The conditions of sulfating roasts, of removing carbon or silicon from metals, and the formation of slag phosphate are instances of this kind. These reactions, too, occur at remarkably low temperatures. Of the alkaline earth oxides, barium oxide reacts at the lowest temperature, followed by strontium oxide and calcium oxide; calcium oxide, like magnesium oxide, generally gives a rapid reaction at about 500°C.

As it was of theoretical interest to ascertain whether alloys of one positive and one acid-forming metal reacted in the same way, and as this question was of practical value in understanding the corroding action of alloys on the linings of crucibles or furnaces, the writer and F. Ilander (38) studied the reactions which occur when calcium oxide is heated with alloys of copper and tin or of copper and antimony in air or oxygen. These "stannides" and "antimonides" react in full analogy with sulfides, carbides, and silicides, forming stannate and antimonate. It is of no consequence whether the tin or antimony were originally present as true compounds with copper, or as solid solutions or even as fused mixtures. The reactions set in very strongly at changes of phase, as from solid to solid, or from solid to liquid state. Thus, for instance, the formation of stannate in these mixtures was found to begin several hundred degrees lower than when heating calcium oxide with stannic oxide, showing a real nascent state effect in the solid state. When homogeneous solid phases decompose or split up into other compounds or solid solutions or in certain cases when a fusion is formed, the matter in question possesses a considerably increased reactivity, in full analogy with the reactivity due to displacements in the lattice as described above. This reactivity can of course be utilized in suitable cases for the production of substances which are otherwise obtained only with difficulty. These conditions are also very clearly revealed in a research by the writer and L. Lindblad on the oxidation of cadmium-silver alloys. When alloys of the approximate composition Ag-Cd (1:1) were used, the formation of cadmium oxide showed a rapid increase at about 425°C., which is in accordance with the results of the work of Natta and Freri (52) and of Åstrand and Westgren (1), which show a change of phase from a low temperature lattice of cesium chloride type to another built with hexagonal close-packing of the atoms. A series of experiments made by the writer and O. Persson have given similar results with CoSn at 515°C., whereas Co<sub>2</sub>Sn, in accordance with the above-mentioned rule, does not give such a discontinuous change in oxidizability.

## THE MECHANISM OF THE REACTIONS

As soon as sufficient material was available in this new field to make possible a preliminary survey, it of course became a pressing problem to find an answer to questions concerning the kinetics and equilibria in these reactions. A question of the greatest interest was to ascertain the function of the lattice ions. As for the problem of the equilibria this was already solved in principle through the researches of van't Hoff (65). Exothermic reactions between solely solid phases should go to completion if solid solutions are not formed, and when the process of repeated grinding and heating was resorted to, some of the early experiments mentioned above showed this to be the case. These problems have been theoretically treated by Tammann (58) with special regard to such reactions as were not known at the time of van't Hoff. On the other hand, reactions leading to solid solutions, e.g.,  $\text{Ag}_2\text{S} + 2 \text{CuI} = \text{Cu}_2\text{S} + 2 \text{AgI}$ , where the iodides or sulfides give solid solutions, were investigated a few years ago by Tubandt (63), who found evidence that in such systems equilibria are obtained which as usual are dependent on the temperature. After some preliminary researches by Tammann on the kinetics of reactions between solids, these problems have been taken up in connection with investigations of the equilibria in endothermic reactions, especially by W. Jander (48). Tammann had shown that after the formation of a layer of reaction product, the reactions proceed, as could be expected because one or more of the reacting particles or groups diffuses through this layer, the growth of which could be calculated from an empirical formula. In these experiments (61) with tungsten trioxide and molybdenum trioxide, that migration of acid anhydride groups which had already been adopted as the explanation in our first attempt at an interpretation of these reactions, was proved. Jander has later proceeded further and has derived from the theory of diffusion an expression for the velocity of reaction in isothermal processes which agrees very well with the observed results. He has also succeeded in calculating the change of reaction velocity with temperature, and in diffusion experiments he has directly proved that the acid anhydrides in fact possess the ability to migrate (49).

There is consequently no doubt that reactions between oxides and salts of oxy acids, (i.e., those first found and described here, such as  $\text{CaO} + \text{CuSO}_4 = \text{CaSO}_4 + \text{CuO}$ ), can proceed exactly according to our first hypothesis, viz., by a migration of the anhydride corresponding to the acid (sulfur trioxide in the above reaction) to the added oxide. But it is of course (as mentioned by Jander) not impossible that reactions between solid substances in other cases can be ascribed to moving metallic oxides or

that they can be reactions between ions. For such processes and for the addition of oxides of alkaline earth metals, the probability of a reaction caused by moving ions is very small even if the salt in question consists of an ionic lattice at ordinary temperatures, for the inability of these oxides to transmit electric current by moving ions (46), even at very high temperatures, shows that the energy required to remove an ion is too large to permit the assumption of ionic mobility which is necessary for this kind of reaction (55). For other oxides the proportion between the energies required for removing an ion or an electron may of course be entirely different, and it must be determined in each special case from such calculations, determinations of conductivity or general chemical considerations, whether the reactions are of ionic or molecular type. Some work of this kind has lately been done by W. Jander (50). This seems to prove, in good agreement with what has been said before, that the type of reaction changes from one mixture to another according to the conditions specified. For reactions between magnesium oxide and tungstates of zinc, cadmium, nickel, and manganese there is no doubt that it is the  $\text{WO}_3$  group which moves over to magnesium oxide. Determinations of electrical conductivity show electronic transmission within the range of experimental temperatures, and only in such cases can determinations of this kind be considered final. As has been lately shown (29) and confirmed by Jander (51), an ionic conductivity is no criterion that there is a reaction between ions.

That such is the case was shown with a high degree of probability in the first experiments with reactions between oxides of alkaline earth metals and halides, although they seem to show that a close connection could exist between ionic conductivity and yield of reaction. By dissolving the unreacted halide in organic solvents it was possible to make yield-temperature diagrams (30), which for a reaction such as  $\text{BaO} + 2\text{CuX} = \text{BaX}_2 + \text{CuO}$  (where  $\text{X} = \text{Cl, Br, I}$ ) are illustrated by figure 8. If the temperature ranges for the rapid rise of the yield curves are compared with those temperatures at which Tubandt (62) found an appreciable increase of ionic conductivity, there is an obvious agreement which cannot be accidental but must imply a real connection, and which has also been observed in other reactions of this kind (29).

Taking into consideration, however, that the ionic conductivity in these salts must be ascribed to the moving copper ions, it is at least evident that it cannot very well be a possible attack of  $\text{Cu}^+$  on barium oxide that is to be considered as the primary cause of the reaction. A contact between  $\text{Cu}^+$  on one hand and  $\text{BaO}$ ,  $\text{Ba}^{++}$ , or  $\text{O}^{--}$  on the other, can hardly matter from the point of view of reactivity, nor can this be the case with the contact between  $\text{Cl}$  or  $\text{Cl}^-$  with  $\text{O}^{--}$ . More important could be the collision

between  $\text{Cl}^-$  and  $\text{Ba}^{++}$ ; but it can hardly be considered as specially probable merely because  $\text{Cu}^+$  starts moving. Taking into account also the fact that reactions between barium oxide or other oxides of alkaline earth metals and a number of halides, e.g., of copper, silver, lead, and nickel, yield the same chemical result, and that these also occur (though with less intensity) at temperatures where there exists only electronic conductivity, it seems natural to ascribe less primary importance to the mobility of the ions. At least it seems more probable that the connection between increase in yield and increase in ionic conductivity is due to the fact that when at a certain temperature the mobility of the ion rapidly increases, this induces a weakening in the lattice and provides a greater possibility for the exchange of atoms. Such a lattice ought consequently to conduct

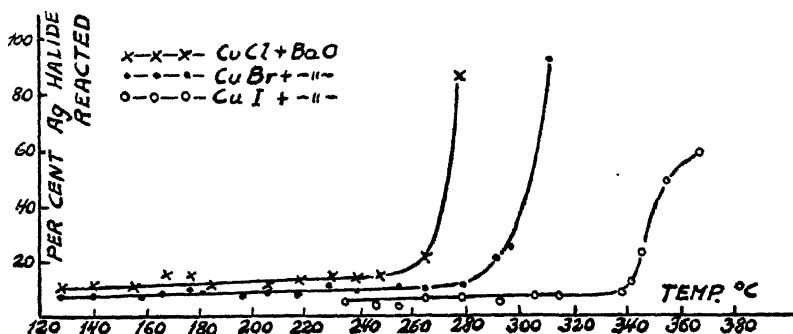


FIG. 8. REACTIVITY OF IONIC LATTICES

itself with regard to reactivity with another in about the same way as during a change of phase.

If we want to test directly the function of the ion in reactions in the solid state we meet with numerous difficulties. First, we must of course distinguish between the influence of the degree of ionization and of the mobility of the ions, and in comparative experiments with different substances must keep one of these factors constant. In preliminary work this has been found extremely difficult with less complicated compounds such as must be used in this case. No doubt it would be simplest to investigate the effect of the mobility of the ions by starting with mixtures of typical ionic lattices and ascertaining whether two such salts, having markedly different conductivity, give yields which are proportionate to their conductivities when reacting with a third substance. To escape complications from difficult diffusion through the reacting layer, the third reactant would probably have to be a gas or possibly a liquid. Still more interest-



ing would be to test the influence of the degree of ionization (by comparing the reactivity with a third substance) in lattices of the ionic type and in lattices where the molecular type is dominant. Here we meet the difficulty, however, of finding a pair of substances with this difference in structure, but with the same internal diffusion. It is also difficult to be sure that the replacement of particles used up by the reaction proceeds with the same velocity in both reacting lattices.

This paper is an attempt to give a survey of the results and problems in a rather new field of chemical research, one which has been allowed to lie neglected for a surprisingly long time although it involves problems in one of the three states of matter, which no one can reasonably allege to be less important or less central than those which are met in the gaseous or the liquid state. In the introduction an explanation has been given for this unfavorable position of the chemistry of the solid state, but it must be remembered that the work in this field is now greatly facilitated by those fundamental earlier researches which apply also to this field, especially those which deal with the processes and laws of heterogeneous systems. It can hardly be denied that the piecing together of the results so far obtained gives the outline of the picture which eventually will appear, even though now it appears as a mosaic with many missing pieces. These results show also that purely chemical methods and reactions are convenient aids in many cases and give important evidence in investigations of crystalline materials and their changes in response to exterior or interior processes.

A large part of this work was made possible through grants from the Research Fund of the Chalmers Institute of Technology or from the Swartz Fund for Researches on Silica.

#### REFERENCES

- (1) ÅSTRAND, A., AND WESTGREN, A.: *Z. anorg. allgem. Chem.* **175**, 90 (1928).
- (2) BALAREW, D.: *Z. anorg. allgem. Chem.* **134**, 117 (1924); **136**, 216 (1924); **160**, 92 (1927).
- (3) BILTZ, W.: *Naturwissenschaften* **13**, 500 (1925).
- (4) BOTHE, W.: *Z. Physik* **20**, 253 (1923).
- (5) BRAGG, W. L.: *Proc. Roy. Soc. London* **114A**, 450 (1927).
- (6) COBB, J. W.: *J. Soc. Chem. Ind.* **1910**, 69.
- (7) COHEN, E., AND HELDERMANN, W.: *Z. physik. Chem.* **87A**, 419 (1914).
- (8) COHEN, E., AND MOESVELD, A.: *Z. physik. Chem.* **85A**, 419 (1913).
- (9) CENTNERSZWER, M., AND AWERBUCH, A.: *Z. physik. Chem.* **115**, 29 (1925).
- (10) DYCKERHOFF, W.: *Dissertation*, Frankfurt a. M., 1925.
- (11) EDNER, A., AND SCHÖNFELD, H.: *Z. Physik* **75**, 442 (1932).
- (12) ENDELL, K.: *Cf. survey in Metall u. Erz* **18**, 1 (1921).
- (13) FAJANS, K.: *Z. Krist.* **61**, 39 (1925); **66**, 321 (1928).

- (14) FENNER, C. N.: *Z. anorg. allgem. Chem.* **85**, 133 (1914).
- (15) GLOCKER, R., AND RISSE: *Z. Physik* **48**, 845 (1928).
- (16) GOETZ, A.: *Proc. Nat. Acad. Sci.* **16**, 99 (1930).
- (17) GOETZ, A., AND HERGENROTHER, R.: *Phys. Rev.* [2] **40**, 643 (1932).
- (18) GOLDSCHMIDT, V. M.: *Ber.* **60**, 1270 (1927).
- (19) HEDVALL, J. A.: Dissertation, Uppsala, 1915; *Z. anorg. allgem. Chem.* **96**, 71 (1916); **98**, 47 (1916); **103**, 249 (1918); **116**, 137 (1921).
- (20) HEDVALL, J. A.: *Z. anorg. allgem. Chem.* **96**, 64 (1916).
- (21) HEDVALL, J. A.: *Z. anorg. allgem. Chem.* **98**, 57 (1916).
- (22) HEDVALL, J. A.: *Z. anorg. allgem. Chem.* **96**, 64, 71 (1916); **98**, 47 (1916); *Z. physik. Chem.* **123A**, 33 (1926).
- (23) HEDVALL, J. A.: *Svensk Kem. Tid.* **1919**, 169.
- (24) HEDVALL, J. A.: *Z. anorg. allgem. Chem.* **120**, 327 (1922); **121**, 217 (1922).
- (25) HEDVALL, J. A.: Survey of previous work in *Svensk Kem. Tid.* **40**, 65 (1928).
- (26) HEDVALL, J. A., AND ANDERSSON, W.: *Z. anorg. allgem. Chem.* **193**, 29 (1930).
- (27) HEDVALL, J. A., AND ASKERLÖF, E.: *Tekn. Samf:s Handl. Göteborg 1929, February*, and *Svensk Gasverksfören:s Årsskr.* **1932**, I, 7.
- (28) HEDVALL, J. A., AND BERGSTRAND, I.: *Z. anorg. allgem. Chem.* **205**, 251 (1932).
- (28a) HEDVALL, J. A., FLOBERG, A., AND PAULSON, P.: *Z. physik. Chem.* In Press (1934).
- (29) HEDVALL, J. A., GARPING, E., LINDEKRANTZ, N., AND NELSON, L.: *Z. anorg. allgem. Chem.* **197**, 399 (1931).
- (30) HEDVALL, J. A., AND GUSTAFSSON, E.: *Svensk Kem. Tid.* **39**, 280 (1927); *Z. anorg. allgem. Chem.* **170**, 71, 80 (1928).
- (30a) HEDVALL, J. A., AND GUSTAFSSON, E.: *Svensk Kem. Tid.* In Press (1934).
- (30b) HEDVALL, J. A., HEDIN, R., AND LJUNGQVIST, S.: *Z. Elektrochem.* In Press (1934).
- (31) HEDVALL, J. A., AND HEUBERGER, J.: *Z. anorg. allgem. Chem.* **122**, 182 (1922); see also **128**, 11 (1923).
- (32) HEDVALL, J. A., AND HEUBERGER, J.: *Z. anorg. allgem. Chem.* **128**, 4 (1923); **135**, 49 (1924).
- (33) HEDVALL, J. A., AND HEUBERGER, J.: *Z. anorg. allgem. Chem.* **128**, 6 (1923); **135**, 65-9 (1924); **140**, 243 (1924).
- (34) HEDVALL, J. A., AND HEUBERGER, J.: *Z. anorg. allgem. Chem.* **135**, 65 (1924).
- (35) HEDVALL, J. A., AND HEUBERGER, J.: *Z. anorg. allgem. Chem.* **140**, 247 (1924); *J. Phys. Chem.* **28**, 1316 (1924).
- (36) HEDVALL, J. A., AND HEUBERGER, J.: *Z. anorg. allgem. Chem.* **140**, 243 (1924). Also HEDVALL, J. A.: *Z. anorg. allgem. Chem.* **162**, 110 (1927).
- (37) HEDVALL, J. A., HEUBERGER, J., AND GUSTAFSSON, E.: *Z. anorg. allgem. Chem.* **135**, 70 (1924); **170**, 80 (1928).
- (38) HEDVALL, J. A., AND ILANDER, F.: *Z. anorg. allgem. Chem.* **203**, 373 (1932).
- (39) HEDVALL, J. A., AND LINDEKRANTZ, N.: *Z. anorg. allgem. Chem.* **197**, 414 (1931).
- (40) HEDVALL, J. A., AND NELSON, L.: *Z. anorg. allgem. Chem.* **197**, 419 (1931).
- (41) HEDVALL, J. A., AND NORSTRÖM, E.: *Z. anorg. allgem. Chem.* **154**, 12 (1926).
- (42) HEDVALL, J. A., AND NORSTRÖM, E.: *Z. anorg. allgem. Chem.* **154**, 1 (1926); *Svensk Kem. Tid.* **1925**, 166.
- (43) HEDVALL, J. A., AND SÖMAN, P.: *Z. Elektrochem.* **37**, 130 (1931).
- (44) HOLGERSSON, S.: Dissertation, Lund, 1927.
- (45) HÜTTIG, G., AND BRÜLL, J.: *Ber.* **65**, 1795 (1932).
- (46) HORTON, F.: *Phil. Mag.* [6] **11**, 505 (1906).

- (47) JANDER, W.: Dissertation, Würzburg, 1927.
- (48) JANDER, W.: Dissertation, Würzburg, 1927; *Z. anorg. allgem. Chem.* **163**, 1 (1927); **166**, 31 (1927); **168**, 113 (1928); survey in *Z. angew. Chem.* **41**, 73 (1928). Cf. also FISCHBECK, K.: *Z. anorg. allgem. Chem.* **165**, 46 (1927).
- (49) JANDER, W.: *Z. anorg. allgem. Chem.* **190**, 65, 397 (1930); **191**, 171 (1930); **192**, 286, 295 (1930); **196**, 321 (1931).
- (50) JANDER, W.: *Z. angew. Chem.* **43**, 1057 (1930).
- (51) JANDER, W., AND STAMM, W.: *Z. anorg. allgem. Chem.* **199**, 180 (1931).
- (52) NATTA, G., AND FRERI, M.: *Rend. accad. Lincei [6a]* **3**, 422 (1927).
- (53) NATTA, G., AND PASSERINI, L.: *Gazz. chim. ital.* **59**, 144, 280, 620 (1929).
- (54) SAITO, H.: *Science Repts. Tohoku Imp. Univ.* **16**, 186 (1927).
- (55) SNEKAL, A.: *Z. tech. Physik* **8**, 577 (1927).
- (56) SCHÄFER, CL., AND SCHUBERT, M.: *Wied. Ann.* **1916**, 283.
- (57) TAMMANN, G.: *Z. anorg. allgem. Chem.* **149**, 53 (1925).
- (58) TAMMANN, G.: *Z. anorg. allgem. Chem.* **149**, 21, 89 (1925).
- (59) TAMMANN, G.: *Z. anorg. allgem. Chem.* **160**, 101 (1927).
- (60) TAMMANN, G., AND WESTERHOLD, FR.: *Z. anorg. allgem. Chem.* **149**, 31, 40 (1925).
- (61) TAMMANN, G., AND WESTERHOLD, FR.: *Z. anorg. allgem. Chem.* **149**, 35 (1925).
- (62) TUBANDT, C., JOST, W., RINDTORFF, E.: *Z. anorg. allgem. Chem.* **165**, 212 (1927); and private communication to the author concerning cuprous bromide and cuprous chloride.
- (63) TUBANDT, C., AND REINHOLD, H.: *Z. physik. Chem.* **140A**, 291 (1929).
- (64) VAN'T HOFF, J. H.: *Z. physik. Chem.* **5**, 322 (1890).
- (65) VAN'T HOFF, J. H., AND VAN DEVENTER, CH.: *Z. physik. Chem.* **1A**, 166 (1887); **5**, 322 (1890).
- (66) WYCKOFF, R. W. G.: *J. Am. Chem. Soc.* **47**, 2866 (1925).
- (67) ŻEMCZUŻNY, S. F., AND BELYNSKY, S. W.: *Z. anorg. Chem.* **59**, 365 (1920).
- (68) ZWICKY, F.: *Proc. Nat. Acad. Sci.* **15**, 253 (1929); **17**, 524 (1931); *Helv. Chim. Acta* **3**, 269, 446 (1930); **6**, 49 (1931).

# HYDRAZOIC ACID AND ITS INORGANIC DERIVATIVES<sup>1</sup>

L. F. AUDRIETH

*Department of Chemistry, University of Illinois, Urbana, Illinois*

*Received August 2, 1934*

## CONTENTS

I. Introduction.....	169
II. Synthesis of hydrazoic acid and azides.....	170
III. Structure of the azide group.....	175
IV. Anhydrous hydrogen azide.....	180
V. Hydrazoic acid (aqueous hydrogen azide).....	182
A. Preparation.....	182
B. General properties.....	183
C. Electrolysis of solutions of hydrazoic acid and its compounds.....	185
D. Reaction with acids.....	186
E. Oxidation.....	187
F. Reduction.....	190
G. Behavior as a nitridizing agent.....	192
H. Detection and determination.....	194
VI. Azidodithiocarbonic acid and azidocarbondisulfide.....	196
VII. Metallic azides.....	198
A. Preparation.....	198
B. General properties.....	201
C. Photochemical decomposition.....	204
D. Thermal decomposition.....	205
E. Explosive properties of the azides.....	207
VIII. Non-metallic azides.....	215
IX. Acknowledgments.....	217
X. References.....	217

## I. INTRODUCTION<sup>2</sup>

From whatever point of view hydrazoic acid be considered, it is bound to excite and stimulate curiosity, if not wonderment and amazement. It

<sup>1</sup> This article is respectfully dedicated to Professor Arthur Wesley Browne, in appreciation of his intensive interest and his extensive investigations in the field of hydrazoic acid and its derivatives and on the occasion of completion by him of twenty-five years of service as Full Professor of Inorganic Chemistry at Cornell University.

<sup>2</sup> It is to be regretted that so much confusion exists in the literature with respect to the nomenclature of  $\text{HN}_3$  and its derivatives. Curtius used the terms "azoimide" and "Stickstoffwasserstoffsäure" (German) for the parent substance. For the organic derivatives the term "azide" and the prefix "azido-" have come into general

contains more nitrogen than any other substance with the exception of nitrogen itself. The ratio of one hydrogen to three nitrogen atoms seems to upset elementary definitions of valence. Its structure has been the subject of considerable discussion and investigation. It has been termed an *ammononitric acid* and a *hydrazonitrous acid* in its relationship to such "parent solvents" as ammonia and hydrazine. As a hydronitrogen it is one of that unique class of hydrogen and nitrogen compounds comparable with the hydrocarbons. Some of its derivatives have received considerable attention because of their explosive character. In some of its reactions it decomposes in such a way that the transitory existence of the free imine ( $\text{>NH}$ ) radical must be assumed. The azide radical has been shown to be a "halogenoid," that is, a halogen-like group, and attempts have been made to prepare free "azine," or "nitrine,"  $(\text{N}_3)_2$ . It is no wonder, therefore, that the literature dealing with hydrazoic acid and its derivatives has grown so extensively and that a real need exists for its compilation. Every attempt has been made to include in this article all important and pertinent facts concerning the chemistry of hydrazoic acid and its inorganic derivatives in as brief and concise a manner as possible.

## II. SYNTHESIS OF HYDRAZOIC ACID AND AZIDES

The first compound containing the  $\text{N}_3$  group was prepared by Peter Griess (113), who obtained phenylazide (diazobenzolimide) by treatment of benzenediazonium perbromide with aqueous ammonia. Emil Fischer (97) also effected the preparation of phenylazide in 1878 by the action of potassium hydroxide or alcoholic hydrogen chloride upon phenylnitrosohydrazine. The very nature of the azide radical with its three nitrogen atoms seemed to preclude the possibility of isolating the parent compound. Its preparation was nevertheless accomplished in 1890 by Theodore Curtius (61, 62, 65) who, by treatment of benzoylazomide (benzoylazide) with sodium hydroxide or sodium ethylate, obtained the sodium salt (Stickstoffnatrium) and from the latter prepared the acid in aqueous solution. Curtius and Radenhausen (71) also prepared pure hydrogen azide, but aban-

---

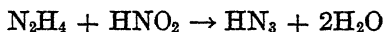
use, although the prefix "triazio-" was also employed by some of the earlier investigators. The latter connotation could not but be ambiguous and soon fell into disfavor. Many American investigators, particularly those interested in studying the inorganic derivatives and reactions of  $\text{HN}_3$ , employ the terms "hydronitric acid" and "trinitrides." English and foreign chemists favor the expressions "hydrazoic acid" and "azides" and these terms are used preferentially by Chemical Abstracts (private communication from Dr. E. J. Crane). The author, in spite of his preference for the names hydronitric acid and trinitrides, has deemed it advisable for the sake of uniformity to use the terminology adopted by the Editorial Board of Chemical Abstracts.

doned its investigation when the latter was seriously injured as the result of an explosion. It was not until 1907 that Dennis and Isham (84) isolated it in larger quantities and studied some of its physical and chemical properties.

All methods which have thus far been employed for the synthesis of hydrazoic acid and the azide group may be divided into five general classes. (These do not include the preparation of azides where the free acid, or one of its derivatives, is used.) (1) Direct synthesis. (2) Action of inorganic and organic derivatives of nitrous acid, upon hydrazine and its compounds involving diazotization of hydrazine or the "hydrazinolysis of nitrous acid." (3) Oxidation of hydrazine. (4) Interaction of certain amides, such as sodium amide, with nitrous oxide or a nitrate. (5) Oxidation of triazenes and the decomposition of higher hydronitrogens.

(1) *Direct synthesis* (160, 249). It has been shown that activated nitrogen reacts with cesium, rubidium, potassium, and sodium to form azides. Except in the case of sodium, nitride formation also takes place presumably by decomposition of the azide which is regarded as the primary product. It is interesting in this connection to note that many years before this work was reported Wendt and Grubb (273a) regarded as highly probable the existence of an ozone analog of nitrogen in the activated gas. They did not mention its possible identity with the free azide radical.

(2) *Interaction of hydrazine and nitrous acid*. This method may be represented in its simplest aspects by the general equation:

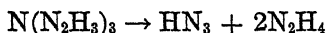


Thus, Thiele (229) found that sodium azide could be obtained in 70 per cent yield by interaction of hydrazine sulfate or hydrate with ethyl nitrite in the presence of sodium hydroxide or sodium ethylate (see also 255). Stolle (215, 216) and others (166, 255) improved the yield by using amyl nitrite in place of ethyl nitrite. Angeli (2) attempted to diazotize both amino groups of hydrazine by allowing silver nitrite to react with hydrazine sulfate, but obtained instead silver azide. Hodgkinson (135) subsequently improved this method for the direct preparation of silver azide. If solutions containing hydrazine sulfate, sodium nitrite, and an excess of sodium hydroxide are refluxed, sodium azide is formed (229). Hydrazine hydrate and sodium nitrite, when heated together in a closed tube at 150°C., give some sodium azide (229). Hydrazine sulfate and potassium nitrite (86, 134) react to give detectable quantities of hydrazoic acid. Nitrosyl chloride (238) reacts with phenylhydrazine to give phenylazide. Nitrogen trichloride (224) and the "red gases" (63) obtained by the action of nitric acid upon arsenic trioxide both react with hydrazine to give hydrazoic acid.

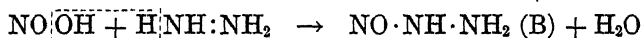
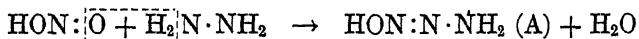
The now classical experiments of Curtius (61), which led to the isolation

of hydrazoic acid, involved the diazotization of benzoylhydrazine to give the corresponding azide. This same method was subsequently employed to synthesize numerous acyl azides ( $RCON_3$ ). The reaction of nitrosamines (210) with hydrazine hydrate, or with a hydrazine salt in the presence of sodium hydroxide or sodium carbonate, yields sodium azide. Certain monosubstituted hydrazines (97) react with nitrous acid to give the intermediate nitrosohydrazines, which in acid or alkaline solution immediately yield the azides.

Whereas Franklin (99, 100, 102, 103) and Turrentine (242) present a great deal of experimental evidence to the effect that hydrazoic acid may be looked upon as an ammonitric acid, Browne (53) concludes that "hydrazoic acid must (also) be in some way structurally related to nitrous acid . . . since . . . all of the above methods involve (the action upon nitrous acid) of some compound or derivative of hydrazine . . . a substance shown . . . to be a dissociating solvent comparable in scope and character with water and ammonia and like these substances constituting the basis of a system of acids, bases, and salts. Hydrazoic acid may therefore be looked upon as *hydrazonitrous acid* . . . experimentally obtainable by hydrazinolysis of nitrous acid and derivable by dehydrazination of normal hydrazonitrous acid."



The partial solvolytic action (5) of hydrazine upon nitrous acid leads theoretically to at least two mixed *aquohydrazonitrous acids*, at least one of which is known in the form of its organic derivatives.

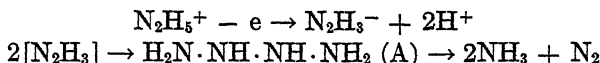


Although no derivatives of 1-hydroxytriazene (A) are known, those of nitrosohydrazine (B) form a large and interesting group. These substances break down in acid and in alkaline solution to yield the azides, whereas thermal decomposition in inert solvents results in the formation of the corresponding amines and nitrous oxide.

The reactions outlined above have been interpreted by Browne as involving initial solvolytic action of hydrazine or one of its derivatives upon nitrous acid or nitrites to give intermediate products which may formally be regarded as derivatives of the hydrazine system and which subsequently undergo desolvation to give hydrazoic acid or azides.

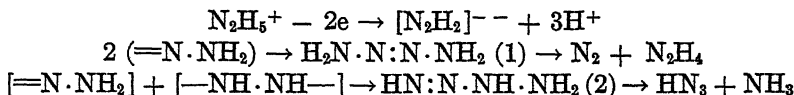
(8) *Oxidation of hydrazine.* Depending upon the oxidizing agent

(delectronator)<sup>3</sup> employed, a variety of products may be obtained from the oxidation of hydrazine (142, 47) in acid solution. Thus, monodelectronators, those active oxidizing agents which accept only one electron per active unit, act upon hydrazine to yield only nitrogen and ammonia. Here the action may be regarded as involving initial formation of the hydrazide radical, two of which subsequently unite to form the intermediate tetrazane (A), which in turn undergoes decomposition to give nitrogen and ammonia.



Such oxidizing agents as the ferric, cupric, nickelic, cobaltic (51), manganic (74), and ceric (15) ions belong to this class.

Didelectronators may bring about complete oxidation of hydrazine to give nitrogen and water, but where such action is incomplete hydrazoic acid is formed along with ammonia. Here the action is believed to involve initial formation of  $[\text{N}_2\text{H}_2]$  radicals to which may be assigned either the symmetrical  $[-\text{NH} \cdot \text{NH}-]$  structure, or the unsymmetrical  $[=\text{N} \cdot \text{NH}_2]$  form. In any case condensation may conceivably take place with formation of either (1) tetrazene, which always decomposes to give nitrogen, or (2) isotetrazene (buzylene), whose organic derivatives decompose to give azides and amines.



Thus, such didelectronators (oxidizing agents which accept two electrons per active unit) as the peroxide (35, 47, 50, 155), persulfate (47, 50), perchlorate, bromate (49), iodate (49, 123), arsenate, antimonate, selenate, and molybdate (51) ions oxidize hydrazine under suitable conditions to give water, nitrogen, hydrazoic acid, and ammonia. The reaction of hydrogen peroxide with hydrazine sulfate has been employed by Browne (35) and by Martin (155) for the preparation of very pure hydrazoic acid.

Kirk and Browne (142) have pointed out that complex delectronators such as potassium permanganate (50, 74) and ammonium metavanadate (48), which undergo reduction in two or more stages, give products indicative of both types of reactions. These oxidants give small quantities of hydrazoic acid when allowed to react with hydrazine in acid solution.

<sup>3</sup> The term "delectronator" has been proposed by Kirk and Browne (142) to cover all oxidizing agents where the action is due to the removal of electrons. Oxidation, augmentation, and delectronation are synonymous terms.



(4) *Ammonolysis of nitrous oxide and sodium nitrate.* Fused sodium amide reacts with nitrous oxide in accordance with the equation:



This reaction was first discovered by Wislicenus (257), but was modified and improved by Dennis and Browne (80) and others (156, 157, 223, 138). It has become the basis of the modern method for the manufacture of sodium azide. This reaction is of particular interest since it represents a clear-cut case of ammonolysis and is further evidence of the fact that "all methods that have hitherto been used for . . . hydrazoic acid may be considered to involve initial solvolysis of an oxyacid of nitrogen or one of its derivatives with potential formation of a solvo-acid belonging to the ammonia system, to the hydrazine system, or to both, and subsequent desolvation of this acid" (53). The substitution of the imide group for its aquo-equivalent, that is, oxygen, in nitrous oxide is strictly ammonolytic in nature. Related to this synthesis is the interesting fact that nitrous oxide, when passed into a liquid ammonia solution of potassium, yields potassium azide (139).

Stolle (217) has extended this method to other amide derivatives and has obtained sodium azide by treating disodium cyanamide with nitrous oxide at temperatures between 300–400°C.

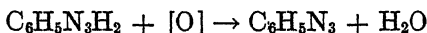
That hydrazoic acid is indeed an ammononitric acid, the ammonia analog of nitric acid, is shown by the fact that sodium nitrate and sodium amide react not only in liquid ammonia (103), but also in the fused state (53), to give sodium azide in accordance with the equation:



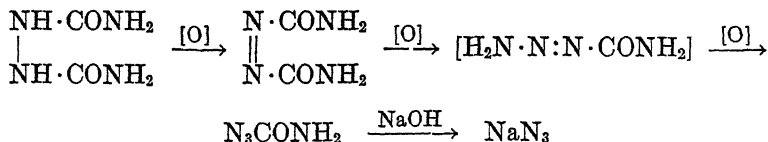
An 80 per cent yield of lead azide can be obtained by the action of an excess of potassium amide upon lead nitrate in liquid ammonia (100). These reactions have been interpreted as involving the conversion of an aquonitrate into an ammononitrate.

(5) *Oxidation of triazenes and the decomposition of higher hydronitrogens* (6). A large number of apparently unrelated reactions may be grouped together under this heading. Most of these are distinctly organic in nature.

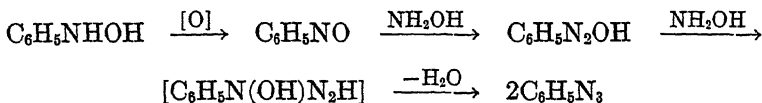
Thus, phenyltriazene in ethereal solution is oxidized to phenylazide by means of sodium hypobromite (88).



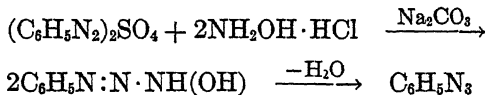
Darapsky (75) obtained sodium azide by oxidation of hydrazidicarbamide with sodium hypochlorite. This method presumably involves the formation of an intermediate triazene derivative.



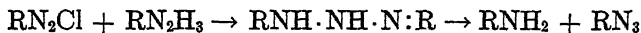
Bamberger (10) obtained phenylazide by treatment of  $\beta$ -phenylhydroxylamine with hydroxylamine in the presence of oxygen. Here again the formation of a hydroxytriazene derivative is assumed, which then breaks down to give the azide and water.



Hydroxylamine salts and diazonium salts (97) react in the presence of sodium carbonate to give the respective azides by way of the intermediate hydroxytriazenes.

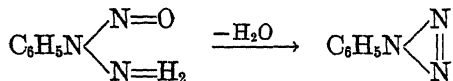


Diazo compounds and hydrazine or hydrazine derivatives react to give the intermediate isotetrazenes (buzylenes) which, under appropriate conditions, break down as indicated by the equation (64):



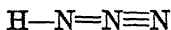
### III. STRUCTURE OF HYDROGEN AZIDE AND THE AZIDES

The structure of hydrazoic acid and its various derivatives has been the subject of considerable discussion and dispute ever since the discovery of the first representative of this class of compounds by Griess (113) in 1866. Emil Fischer (97) prepared phenylazide by the action of dilute alkalis upon the corresponding nitrosohydrazine

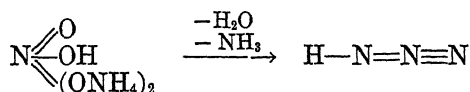


and on the basis of this reaction promulgated the so-called cyclic or ring structure for the azide group. This structure was subsequently accepted by Curtius (61), the discoverer of the parent hydronitrogen.

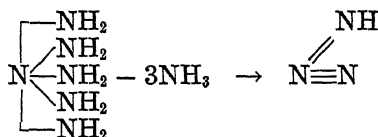
More detailed investigation of the properties of hydrazoic acid and its derivatives soon caused a decided shift in opinion to the chain structure.



Turrentine (242) was one of the first to advocate such a chain structure on the basis of similarity between nitric and hydrazoic acids. If nitric acid be assumed to possess a pentavalent nitrogen atom, it was regarded as highly probable that one of the nitrogens in hydrazoic acid must also be pentavalent. Mendelejeff (158) considered hydrazoic acid as derived from a secondary ammonium orthonitrate, and here again desolvation with loss of water and ammonia



could lead only to a chain structure. From a consideration of hydrazoic acid as a desolvation product of an ammonoorthonitric acid, Franklin (99, 100, 102) also came to the same conclusion.



Thiele (231, 1), reasoning from the action of the Grignard reagent upon hydrazoic acid and its organic compounds, by which diazoamino compounds are formed, also decided in favor of the straight chain formula. It was argued that the formation of such open chains from ring structures could take place only if the ring broke, not at the double link, but at a single link. Sidgwick (200) has pointed out, however, that this piece of evidence is hardly conclusive, since the double link between nitrogen atoms is very strong. Any rupture in a ring containing two doubly linked nitrogen atoms most certainly would not occur at that point.

Langmuir (146) was the first to point out that the cyanate and azide groups are isosteric and that many cyanates and azides possess similar physical and chemical properties. Both potassium azide and potassium cyanate crystallize in the tetragonal system and have approximately the same axial ratios.

For  $\text{KNCO}$ ,  $a:c = 1:0.5766$

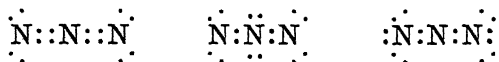
For  $\text{KN}_3$ ,  $a:c = 1:0.5798$

Certainly one might expect isomorphism of cyanates and azides. One might also expect both ions to possess similar structures. Langmuir discards the ring structure in favor of a chain structure which he represents as  $(\text{N}=\text{N}=\text{N})^-$  for the azide ion.

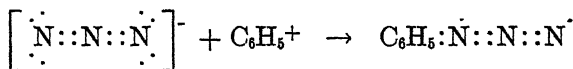
The physical properties of aqueous solutions of the isosteric pairs  $\text{KN}_3$ - $\text{KNCO}$  and  $\text{NaN}_3$ - $\text{NaNCO}$  have been determined by Cranston and Livingstone (58). The values for the refractive indices, solubilities, and equivalent conductivities were found to be so nearly alike that they concluded that the structures of the cyanate and azide ions must also be similar. Günther and Perschke (117, 179) contend, however, that it is hardly permissible to draw conclusions concerning the configuration of ions on the basis of similarity in the physical properties of their solutions. They found that the viscosities, refractive indices, and conductivities of aqueous solutions of sodium azide, thiocyanate, and nitrite were also very much alike, even though, according to them, the structures of these ions are presumably very different.

More recently a general comparison of the crystallographic and physical properties of potassium azide and potassium cyanate also led Erlenmeyer and Leo (94) to assume similarity in structure of the two radicals.

Hendricks and Pauling (130) studied the crystal structure of sodium and potassium azides by x-ray methods and definitely decided in favor of a collinear structure for the azide ion. The positions of the azide ions in the crystal lattices are such that they require centers of symmetry. They therefore discard the idea of a double and a triple link, as promulgated by Turrentine (242), and assign to the azide ion one of the following three electronic structures.

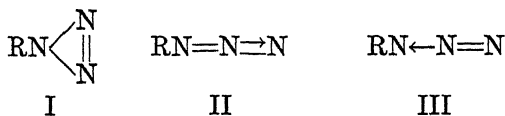


It is now generally conceded that the azide ion possesses a chain structure, but some question still existed until very recently as to those organic compounds in which the azide group is attached by a covalent bond to an organic radical. In a discussion of this problem from the standpoint of the influence of structure upon the dipole moment, Sidgwick (200) says, "We should expect the  $\text{N}_3$  group to have the same structure in the ion as it has in the covalent form. It is to be noticed, however, that the symmetry and non-polarity of the ion upon which its stability may be supposed to depend, cannot be preserved when it becomes covalent; an imaginary phenyl cation can attach itself only to one of the terminal nitrogens, since the central nitrogen has no unshared electrons



and it will then give, if the open chain is maintained, the molecule  $\text{RN}=\text{N} \rightleftharpoons \text{N} \rightleftharpoons \text{N}$  in which the balance is destroyed by the disappearance of one of the

two coördination moments." Three possible structures for organic azides are given by Sidgwick.



Of these I should have a small dipole moment (198, 220) owing to its symmetry, whereas II and III, since they both contain coördinate links, should be characterized by relatively large moments. "A large moment," Sidgwick (200) further states, "would hardly be compatible with the volatility of the azides." The organic azides have boiling points which lie very close to those of the corresponding bromides. Measurement of the dipole moments of a series of organic azides shows that the azide radical has a moment of about 1.5 D., also that the negative end is in a direction opposite from the point of attachment to the organic radical. Structure III is therefore out of the question since its moment is in the wrong direction, whereas II, containing a coördinate link, is not reconcilable with such a low moment as the  $\text{N}_3$  group has been found to possess. The ring structure (222) seems to be most nearly in agreement with the results of these measurements.

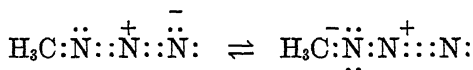
In spite of these considerations Bergman and Schütz (16, 17) feel that the results of dipole measurements are hardly conclusive. Upon analogy with the structure of the mustard oils they are inclined to favor a chain structure even for the organic azides.

The problem has been approached from still another angle by Lindemann (147, 148, 149). Parachor values for a series of organic azides agree very closely with those required for a ring structure. It should be pointed out however that these values differ but slightly from those required for chain structures with coördinate linkages.

Hantzsch (126) investigated the absorption spectra of various azides and as the result of these studies divided all derivatives of hydrazoic acid into two groups. The first group is termed the *ester type*. These are non-electrolytes with discontinuous or selective absorption. They presumably possess an altogether different structure from the *true salts*. The latter class includes the alkali and ammonium azides which are typical electrolytes. Hantzsch is inclined to favor a ring structure especially for the ester type of compound.

However, the most recent work seems again to favor a chain structure for all azides. Sutton (221) studied the crystal structure of cyanuric triazide,  $(\text{CN})_3(\text{N}_3)_3$ , by means of x-rays and assigned to the molecule a

structure in which the azide groups are in the form of short chains attached to the cyanuric ring through the middle nitrogen atom. Brockway and Pauling (34) studied methyl azide,  $\text{CH}_3\text{N}_3$ , by means of the electron diffraction method, and found their results to be compatible only with a state of resonance between two linear structures having the following electronic configurations (see structures II and III suggested by Sidgwick)



Sidgwick, Sutton, and Thomas (201) seem also lately to admit of the possibility of a chain structure for the azide group and state that "dipole moments are compatible only with a ring structure or resonance between the two linear structures." In his recent discussion of this problem before the Faraday Society at Oxford in April, 1934, Sidgwick (199) points out that for such tautomerism to exist and show a reduced moment it is necessary that the time of conversion of one form into the other take place in less than  $10^{-8}$  seconds (the relaxation time for the rotation of the dipole in the field). Sidgwick also discusses the bearing of the data for the heat of formation of the  $\text{N}_3$  group as determined by Roth and Müller (195). These investigators found the heat of formation of the  $\text{N}_3$  group from its atoms to be 210 kg-cal. Now the heats of formation of the various N—N linkages can be calculated and on this basis the expected values for the three structures determined. They are:

I. $2(\text{N}—\text{N}) + (\text{N}=\text{N})$	$= (2 \times 37.6) + 95.3$	$= 170.5$
II. $2(\text{N}—\text{N})$	$= 2 \times 95.3$	$= 190.6$
III. $(\text{N}—\text{N}) + (\text{N}=\text{N})$	$= 35.7 + 208$	$= 245.6$
Observed value		$= 210.5$

The observed value lies about halfway between the two calculated values for the chain structures, but the energies of the two forms as calculated are so different as to make a resonance between them in the Pauling sense impossible—even if the ring structure is definitely excluded by these figures. Sidgwick points out that it is probable that the value assumed for structure III is much too high. By introducing a correction for the  $\text{N}=\text{N}$  linkage, a new value of 181 kg-cal. for structure III is obtained. This would then bring the calculated values for forms II and III very close to each other, thus fulfilling the necessary requisites for such a state of resonance requiring that the energy levels of the two forms be very close to each other. That the actual molecule has a value higher by some 20 kg-cal. might be expected as the result of resonance.

## IV. ANHYDROUS HYDROGEN AZIDE

Pure hydrogen azide is a colorless, extremely explosive liquid, boiling at  $37^{\circ}\text{C}$ . and melting at  $-80^{\circ}\text{C}$ . Dennis and Isham (84) recommend the following method of preparation: Sulfuric acid is allowed to react, drop by drop, with solid potassium azide. The hydrogen azide formed is carried over in a stream of carbon dioxide-free air from the distilling flask through several drying columns containing calcium chloride into a receiver surrounded by an ice-salt mixture. To prevent loss of hydrogen azide, absorbing vessels containing absolute methanol are attached to the receiving vessel. The whole apparatus should be placed behind a heavy screen of 2-inch pine planking, and all manipulations of stopcocks etc. should be effected from the outside. Double windows of half-inch plate glass, suitably placed in such a screen, can be used for purposes of observation. Too much attention can not be called to the extremely treacherous nature of hydrogen azide.

The physical properties of hydrogen azide are none too well defined. Dennis and Isham (84) state that the pure compound is probably heavier than water, exhibits high vapor tension and possesses the normal vapor density twenty-five degrees above its boiling point. Hydrogen azide is soluble in water, the alcohols, and ether. It may be extracted from aqueous solution by ether, as the distribution ratio of hydrogen azide between these two solvents (water: ether) is approximately 1:7.

Like the anhydrous hydrogen halides, so liquid hydrogen azide has been found to act as a solvent for many substances. McKinney (153) investigated qualitatively the solubilities of several hundred of the more common inorganic compounds in liquid hydrogen azide. His method consisted of condensing hydrogen azide upon a small sample of the dry salt contained in a specially designed solubility cell immersed in ice water. The salt was permitted to remain in contact with the liquid for at least thirty minutes, during which time it was observed closely for evidences of physical and chemical action. The solution was then filtered and siphoned over into an electrolytic cell fitted with two platinum electrodes. The filtered liquid was subjected to electrolysis for several minutes. If nothing untoward had happened up to this point the solvent was removed by evaporation and the residual material tested qualitatively. In about ten per cent of the experiments violent explosions occurred during electrolysis of the filtered solutions.

The results of this investigation by McKinney may be summarized as follows:

Acetic acid is more or less completely displaced from ammonium, cadmium, chromium, cobalt, mercuric, nickel, and sodium acetates,

resulting in the formation of the corresponding azides and acetic acid. Ammonium acetate swells to about double its original bulk when hydrogen azide is condensed upon it. The possible formation of a solvate containing hydrogen azide of crystallization is suggested. In several instances none of the metal associated with the acetate was found in the filtered liquid, indicating that the azides of cadmium, lead, copper, and manganese are insoluble in hydrogen azide.

A similar solvolytic action is observed in the case of the arsenites. Thus, calcium, cupric, ferric, and potassium arsenites react with hydrogen azide to give solutions from which arsenic is deposited upon electrolysis. Only in the case of the ferric and potassium arsenites were the cations found in the filtered solution.

The following halides were found to be appreciably soluble in hydrogen azide at 0°C.: ammonium bromide, ammonium iodide, bismuth iodide, cobaltous chloride, magnesium chloride, manganous chloride, mercuric chloride, nickelous chloride, potassium chloride, potassium bromide, potassium iodide, sodium bromide, and sodium iodide. Mercuric chloride is very soluble in hydrogen azide and yields cathodic mercury upon electrolysis. Iodine dissolves in hydrogen azide to give a reddish solution which appears to conduct the electric current fairly well.

The nitrates are in most instances only slightly soluble. Manganous, cobaltous, and nickelous nitrates are moderately soluble. Electrolysis of a solution of bismuth nitrate gave a cathodic deposit of bismuth.

All chromates with the exception of barium chromate were found to react with or dissolve in hydrogen azide. In some instances (lead chromate, mercuric chromate, lithium chromate) a green solution was obtained indicating reduction of chromium to the trivalent state.

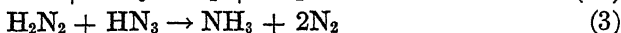
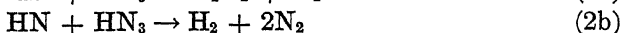
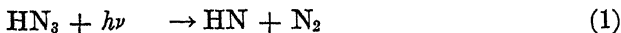
Cadmium, zinc, and potassium iodates react with the solvent to liberate free iodine. Sulfates, oxides, carbonates, and sulfides, in general, do not react with or dissolve in hydrogen azide.

Pure hydrogen azide is a poor conductor. However, it has been shown by Browne and Lundell (46) that the addition of potassium azide results in the formation of a conducting solution. These investigators electrolyzed solutions of potassium azide in hydrogen azide and found the ratio of hydrogen to nitrogen evolved to be approximately 1:3, but somewhat lower under certain conditions. Ammonia was usually found to be present in the residual electrolyte, but no hydrazine. The extreme explosiveness of the residual liquid after electrolysis was attributed to the possible formation of azine (nitrine, nitrazone,  $(N_3)_2$ ).

Ultra-violet light brings about decomposition of gaseous hydrogen azide with formation of nitrogen, hydrogen, and ammonium azide. Light absorption of hydrogen azide sets in at 2200 Å. U. and decomposition



certainly takes place at wave-lengths below 2400 A. U. The mechanism offered tentatively by Beckman and Dickinson (11, 12) is given by the following equations.



At low pressures (4 cm.) hydrogen azide decomposes with measurable rapidity at 290°C. (181). Under these conditions about 9 to 11 per cent of  $\text{HN}_3$  is decomposed in 25 minutes. A white solid, presumably ammonium azide, is formed along with gases non-condensable in liquid air. Thermal decomposition appears to be influenced catalytically by the walls of the Pyrex glass reaction chamber.

The molecular heat of decomposition of pure liquid hydrogen azide into nitrogen and hydrogen at constant pressure has been calculated by Roth and Müller (195) from the heat of combustion of the phenyl azide and ethyl azidoacetate. By subtraction of the heat values for specific groups associated with hydrogen azide in both of these molecules the values 64.3 kg-cal. and 70.6 kg-cal. are obtained. An average value of 67 kg-cal. per mole is assigned as the heat of decomposition of liquid hydrogen azide into nitrogen and hydrogen at constant pressure.

## V. HYDRAZOIC ACID (AQUEOUS HYDROGEN AZIDE)

### A. Preparation

Hydrazoic acid is obtained in solutions of varying concentrations by treatment of aqueous solutions of potassium or sodium azides with dilute sulfuric acid. Because of the high volatility of hydrogen azide, it is advisable to bring the solution of the azide to the boiling point and then add sulfuric acid drop by drop. If this is not done a very concentrated solution of hydrazoic acid comes over at first, which because of its explosibility is apt to be hazardous.

The interaction of sodium azide with aqueous solutions of fluosilicic (138) or oxalic acids (93, 138) has been suggested for the preparation of hydrazoic acid. Pure hydrazoic acid has also been obtained by interaction of barium azide with dilute sulfuric acid (138).

The methods outlined under the oxidation of hydrazine in acid solution result in the formation of hydrazoic acid as an aqueous distillate. Thus, oxidation of hydrazine in the presence of sulfuric acid by hydrogen peroxide gives pure hydrazoic acid (35, 155).

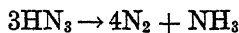
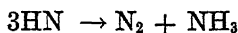
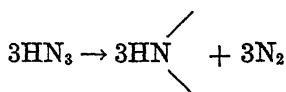
If aqueous solutions of hydrazoic acid are distilled, some hydrogen azide

comes over first in the gaseous form; this is followed at temperatures between 90–100°C. by a very concentrated aqueous distillate, the first fractions of which approximate a 27 per cent solution of  $\text{HN}_3$ . Continued distillation finally causes an equilibrium to set in, after which a very dilute solution is obtained to the last drop. These findings would seem to indicate the existence of a constant-boiling mixture of hydrogen azide and water, but no experimental work to verify this supposition has been reported.

### *B. General properties*

Pure aqueous solutions of hydrazoic acid may be kept indefinitely. The solutions are characterized by an intense, penetrating odor and the vapor when inhaled even in small concentrations induces decidedly unpleasant physiological effects. Headaches and dizziness are produced, with a clogging of the mucous membrane, giving all the effects of a severe head cold. The acid and its salts act as protoplasmic poisons. Subcutaneous injection of small doses of the sodium salt (0.03 g.) into mammals produces spasms and symptoms of heart and lung paralysis (150). In the case of frogs, respiration is arrested, followed by convulsions (20, 21). Growth of plants is inhibited and the germination of seedlings is prevented (197). Several cases of poisoning by hydrazoic acid have been reported (143, 211). It has been suggested that sodium azide be used as an agent for poisoning harmful animals (251).

Platinum black, platinum sponge, and platinum foil (168) decompose solutions of hydrazoic acid with the formation of ammonia and nitrogen. Occluded hydrogen or oxygen accelerate this decomposition. The catalytic decomposition of hydrazoic acid presumably involves the intermediate formation of free "imino" radicals which decompose further to give more nitrogen and ammonia.



Curiously enough, however, colloidal platinum, which might be expected to produce the greatest effect, does not decompose hydrazoic acid under any conditions. All salts of hydrazoic acid (with the exception of  $\text{Fe}(\text{N}_3)_3$ ), as well as the acid itself, inhibit the catalytic decomposition of hydrogen peroxide by colloidal platinum (171, 172, 174). Colloidal palladium does

not affect sodium azide solutions, but in the presence of hydrogen brings about reduction with evolution of nitrogen and ammonia (254).

Hydrazoic acid is decomposed by ultra-violet light (111) to give a variety of products depending upon conditions. In dilute acid solutions hydroxylamine and nitrogen are the main products, along with some ammonia. In alkaline solutions hydrazine is also formed, along with ammonia and hydroxylamine. Primary decomposition to yield "imine" radicals again seems highly probable. (See pp. 187 and 194.)

Hydrazoic acid is a monobasic acid and in many of its properties and reactions shows a striking similarity with the halogen hydracids. Birkenbach and Kellermann (34) have shown that the azide radical lies between chlorine and bromine in the relative order of reactivity of various halogen and halogenoid groups. For that reason hydrazoic acid has been termed a "halogenoid" (248) hydracid. When added to solutions of soluble silver, lead, and mercurous salts the corresponding azides are precipitated. Many of its salts, such as those of aluminum, chromium, and ferric iron, are appreciably hydrolyzed in aqueous solution. When solutions of these azides are heated the corresponding hydroxides are precipitated, indicating that hydrazoic acid is a very weak acid indeed.

These findings have also been checked by conductivity measurements. The results tabulated below, taken from determinations by West (252), indicate that at 25°C. hydrazoic acid is somewhat stronger than acetic acid (71). ( $V$  = dilution in liters per mole,  $\mu$  = molecular conductivity,  $K$  = dissociation constant.)

$V$	$\mu$	$K \times 10^{-5}$
10	5.38	1.98
100	15.98	1.80
1000	45.97	1.66

West determined the conductance of sodium azide at various dilutions and extrapolated these results to infinite dilution obtaining a  $\Lambda_{\infty}$  value for  $\text{NaN}_3 = 109$ . To determine the dissociation constant  $K$  for hydrazoic acid it was necessary to have the limiting value for  $\text{HN}_3$ , which was calculated on the basis of the equation

$$\Lambda_{\infty \text{HN}_3} = \Lambda_{\infty \text{NaN}_3} + l_{\text{H}^+} - l_{\text{Na}^+} = 109 + 320.5 - 44.5 = 385$$

The value for the dissociation constant,  $K$ , was also determined from the rate of inversion of cane sugar, as induced by hydrazoic acid, and was found to be  $1.86 \times 10^{-5}$ . No experimental details are given by West.

Very few other physical constants for hydrazoic acid have been determined. Berthelot and Matignon (18) calculated the heat of formation from the heat of combustion of the ammonium salt.



Roth and Müller (195) recalculated the data and introduced certain corrections giving a value for the heat of formation of hydrazoic acid of  $-53.4$  kg-cal. These figures are quite in accord with the extreme explosiveness of the substance. The heat of solution at  $11^{\circ}\text{C}$ . was found to be  $-7.08$  kg-cal. The heat of neutralization by barium hydroxide was found to be  $10.0$  kg-cal.; by ammonia,  $8.2$  kg-cal. This latter figure agrees well with an independent observation by Bach (9), who obtained the value  $8.3$  kg-cal.

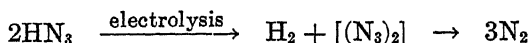
### *C. Electrolysis of solutions of hydrazoic acid and its compounds*

Many attempts have been made to discharge the azide ion directly by electrolysis, both in aqueous and non-aqueous solutions (183). While the presence of free  $(\text{N}_3)_2$  as a product of electrolytic oxidation has never been demonstrated conclusively, it is very probable that the triatomic  $\text{N}_3$  group (245) is the primary electrolytic product. Its existence is only transient, since it breaks down rapidly to give ordinary nitrogen. In all cases electrolysis of aqueous solutions of sodium azide (32, 177, 178, 223), ammonium azide (133), and hydrazoic acid (177, 178, 223), using inert electrodes, results in the discharge of cathodic hydrogen and anodic nitrogen. The ratio of the volumes of hydrogen to nitrogen formed is usually about 1:3, as might be expected theoretically. Ammonia and hydrazine are very often found in the residual electrolyte, indicating that partial reduction has also taken place.

Turrentine (240) electrolyzed 3 per cent sodium azide solutions using various metals as anodes and in the case of magnesium, aluminum, zinc, and cadmium obtained corrosion efficiencies exceeding 100 per cent. He accounted for this by assuming that these metals dissolved at a valence lower than usual and were subsequently oxidized. Ammonia was usually found to be present in the residual electrolyte.

Electrolyses of solutions of hydrogen azide in liquid ammonia (43)—solutions prepared by dissolving ammonium azide in liquid ammonia—give varying results, depending upon the electrode materials employed. When smooth platinum electrodes are used the ratio of  $\text{H}_2:\text{N}_2$  averages 1.93:1.00, instead of 1:3 as might be expected if the azide ion only is discharged. It must therefore be assumed that either the discharged azide ion reacts with the solvent to liberate nitrogen and regenerate ammonium azide (43), or that amide ions (from ammonia as solvent) are also discharged when a platinum anode is used. In either case, the ratio of  $\text{H}_2:\text{N}_2$  would then be 3:1. The experimentally determined ratio is a smaller one, indicating that some azide ions are also discharged. When graphite electrodes

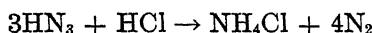
are employed a ratio of  $\text{H}_2:\text{N}_2$  equal to 1:3 is obtained, indicating that under these conditions only the azide ion is discharged.



When other anodes (44) are used in the electrolysis of ammonium azide in liquid ammonia electrolytic corrosion takes place. With copper anodes, cuprous and cupric azides are obtained. With silver, cadmium, lead, and antimony the normal azides,  $\text{AgN}_3$ ,  $\text{CdN}_6$ ,  $\text{PbN}_6$ , and  $\text{SbN}_3$ , are formed without liberation of gas at the anode. Aluminum, iron, and nickel anodes also undergo electrolytic corrosion accompanied by liberation of nitrogen gas and formation of the corresponding ammono-basic azides.

#### *D. Reaction with acids*

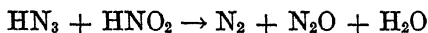
Hydrochloric acid reacts slowly with hydrazoic acid in accordance with the equation (83):



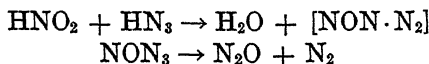
When a mixture of the two acids is boiled some chlorine is also evolved (40). Dry hydrogen azide in contact with dry hydrogen chloride reacts rapidly with formation of solid ammonium chloride (83). It is interesting to note that freshly prepared mixtures (242) of the two acids have the ability to dissolve such noble metals as platinum and gold.

Hydriodic acid (40, 137) reduces hydrazoic acid to give free iodine, nitrogen, and ammonia. Hydrobromic acid is nitrized with liberation of bromine (100).

According to Werner (250) and to Thiele (229), nitrous acid reacts with hydrazoic acid to give nitrogen, nitrous oxide, and water. This reaction has been shown to proceed quantitatively in accordance with the equation (206):



It has been suggested as a method for the determination of nitrites in the presence of nitrates. The reaction is essentially one characteristic of all secondary amines. Thus, it is assumed that the two substances react first to give nitrosyl azide, which then breaks down to give nitrous oxide and nitrogen. The intermediate compound has never been isolated.

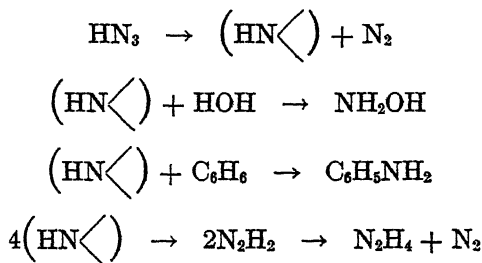


The reaction with sulfuric acid has been the subject of considerable investigation. Hantzsch (125) carried out molecular weight determina-

tions of hydrogen azide in 100 per cent sulfuric acid and concluded that a compound of hydrogen azide with three molecules of sulfuric acid,  $\text{HN}_3 \cdot 3\text{H}_2\text{SO}_4$ , was formed. When hydrogen azide is passed into concentrated sulfuric acid a white crystalline product (83) is obtained which is decomposed by water to give hydrazine sulfate.<sup>4</sup> The original white product has been shown to be hydrazine disulfate (152),  $\text{N}_2\text{H}_4 \cdot 2\text{H}_2\text{SO}_4$ , a substance first prepared by Sommer (208) in an entirely different fashion. Treatment with water gives the normal sulfate and free sulfuric acid.

The interaction of sulfuric acid with hydrazoic acid has also been investigated by Schmidt (196). If these two substances are heated together in a fairly concentrated solution considerable quantities of hydroxylamine and some ammonia, but no hydrazine, are formed. The presence of the sulfite radical in the residual solution indicated partial reduction of the sulfuric acid. If sulfuric acid be permitted to react with a benzene solution of hydrogen azide at room temperatures, gas evolution takes place until all of the hydrogen azide has disappeared. Large quantities of hydrazine sulfate and some hydroxylamine and aniline sulfate form under these conditions. If this decomposition be allowed to take place at 60°C., the main product is aniline sulfate, with little or no hydrazine sulfate.

Schmidt regards these reactions as being due to the primary formation of the "imino" ( $\text{HN} \langle$ ) residue, which may combine with water to form hydroxylamine (see p. 194) or with benzene to form aniline. Two imino residues may also combine to form diimide, which disproportionates itself to give nitrogen and hydrazine.

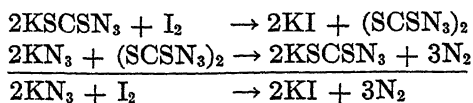


#### *E. Oxidation*

Depending upon conditions a variety of products is obtained by the action of various oxidizing agents upon hydrazoic acid and azides. In

<sup>4</sup> Browne and Wilcoxon (53) have called attention to this fact as a basis for their contention that sulfuric acid may be considered not only a dehydrating agent, but a desolvating agent in general. The reaction with hydrogen azide is an example of "dehydration" by virtue of the fact that it may be interpreted as the removal of hydrazine from the hydrazonitrous acid,  $\text{HN}_3$ .

*neutral* solution iodine reacts with azides in the presence of thiosulfates (185), sulfides (185), azidodithiocarbonates (39), and carbon disulfide (39, 95) to liberate nitrogen. Iodine itself does not affect solutions of azides, but the tetrathionate ion formed by the interaction of iodine with the thiosulfate catalyzes the reaction. In the case of the azidodithiocarbonates it has been shown by Browne and Hoel (39) that the formation of azidocarbondisulfide,  $(SCSN_3)_2$  (see p. 196), catalyzes the reaction. The azidodithiocarbonate and the halogenoid radical undergo repeated mutual conversion into each other, constituting what has been termed a reciprocal catalytic pair. The reaction may be represented by the following series of equations:

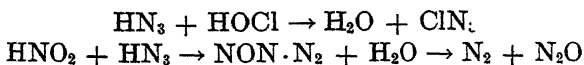


The oxidation of azides by iodine in the presence of carbon disulfide (95) is of course due to the formation of some azidodithiocarbonate, which then catalyzes the reaction as indicated above. These reactions permit the quantitative determination of alkali azides, either gas-volumetrically (185) or by the determination of the excess of standard iodine solution by means of arsenite solutions (95). It should be pointed out in this connection that the action of iodine upon silver azide results in the formation of iodine azide (124).

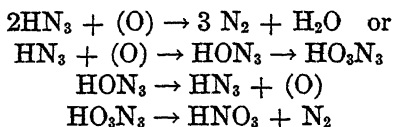
In acid solution such oxidizing agents (194) as ferric salts, iodates, and hydrogen peroxide have no appreciable effect. Potassium chlorate and manganese dioxide react only slightly, while potassium permanganate and persulfates are more vigorous in their action. Complete and rapid reaction is obtained only when the ceric salts (207), hypochlorous acid (184), or nitrous acid (206, 207, 230, 169) are used. Oxidation with complete conversion of the azide group into molecular nitrogen is obtained only in the case of ceric salts, in accordance with the equation:



This reaction is used for the quantitative determination of azides. Either the volume of nitrogen (207) evolved is measured directly, or the excess of standard ceric salt solution determined iodimetrically (155). In the case of hypochlorous acid and of nitrous acid, condensation takes place to give in the first instance chlorazide, and in the second case nitrosylazide (207), which decomposes into nitrogen and nitrous oxide.



Dennis and Browne (80, 81) investigated potassium permanganate as an oxidizing agent in the hope that it might lead to a quantitative method for the determination of hydrazoic acid. They found the relationships to be very complex and were unable to assign any definite stoichiometric ratios to the amounts of permanganate and hydrazoic acid used. The gases evolved contained oxygen in addition to nitrogen. Nitric acid was also found to be present in the residual solution. Raschig (182a) subsequently studied this reaction and found that the addition of potassium iodide to the decolorized solution resulted in the liberation of iodine. The iodine set free was titrated with standard thiosulfate solution, but curiously enough, such a solution would shortly again turn blue, owing to the liberation of more iodine, requiring the further addition of thiosulfate. This seemed to indicate the presence of some oxidizing agent which acted but slowly upon hydriodic acid, possibly a hypoazidous acid ( $\text{HON}_3$ ), the analog of hypochlorous acid. Such oxyacids had already been discussed by Dennis and Browne (80, 81), who argued that the halogenoid character of the azide radical might lead one to expect at least the transient formation of such substances. Raschig assumed that the hypoazidous acid would decompose to give either oxygen and hydrazoic acid, or undergo autooxidation with formation of some of the chlorite and chlorate analogs of hydrazoic acid,  $\text{HO}_2\text{N}_3$  and  $\text{HO}_3\text{N}_3$ . The oxidation of hydrazoic acid by permanganate may therefore be summarized by the following series of equations:

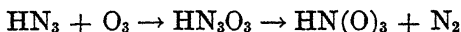


Such a mechanism would account for the complex character of the reaction.

Most oxidizing agents have no effect upon alkaline solutions of azides. However, if a 10 per cent ozone-oxygen mixture is passed into a sodium azide solution, the latter turns yellow, eventually assumes a deep orange color, and finally fades again on long-continued action. Complete ozonation results in the disappearance of all traces of the azide ion. The residual solution becomes more strongly alkaline and possesses powerful oxidizing properties. Acidification immediately destroys any color, whereas increase in alkalinity stabilizes the orange compound formed under these conditions. Gleu and Roell (112) attempted to isolate this colored compound, but were unsuccessful. Heating of the ozonized solutions resulted in the liberation of considerable quantities of oxygen. If any unreacted azide was still present, nitrogen and nitrous oxide were also obtained. The addition of bicarbonate solution resulted in the discharge of the color, but evolution of little oxygen. Nitrite and nitrate were found in such solutions. In many

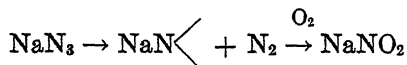


respects these ozonized solutions behaved as though a salt of a per-acid were present. Gleu and Roell assume the direct addition of ozone to the azide with formation of a compound which might be regarded as one of the hypothetical oxyacids mentioned before. This substance is thought to decompose to give nitrogen and pernitrite, that is, a salt of a *pernitrous* acid to which the formula  $\text{H}-\text{O}-\text{ONO}$  is assigned. In acid solution this substance immediately decomposes with formation of some nitric acid.



It might be mentioned in this connection that the direct addition of ozone is not unexpected, particularly since all structures for the azide ion assume the presence of a double bond. Ozone adds very easily to organic compounds containing a double bond.

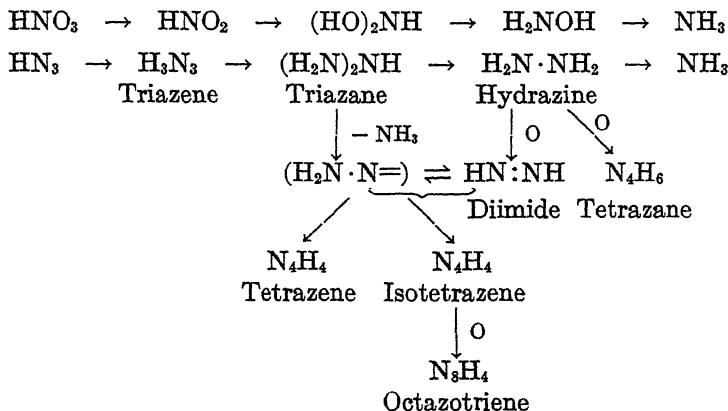
K. A. Hofmann and U. Hofmann (136) studied the action of molecular oxygen on sodium azide-sodium hydroxide melts and observed the formation of sodium nitrite. They assumed the direct addition of oxygen to the imide radical with formation of the nitrite. Gleu and Roell, however, are not inclined



to accept this mechanism for the reaction studied by them, since ozonization of other substances containing the imide grouping does not result in the formation of pernitrous acid.

#### F. Reduction

From a consideration of hydrazoic acid as ammononitric acid it is possible to represent the action of reducing agents diagrammatically (6, 264) as indicated by the following scheme:

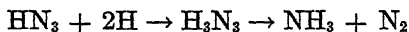


Naturally, the reverse series of reactions points to the possible nitridation of other hydronitrogens to hydrazoic acid. Thus, nitridation of hydrazine to hydrazoic acid is quite analogous to the complete oxidation of hydroxylamine to nitric acid.

The mechanism outlined above is interesting in that it brings out the interrelationship of no less than ten different hydronitrogens (6). It also serves to characterize all of these as ammonio compounds from the Franklin (99) point of view.

Investigators are generally agreed that triazene is the first step in the reduction of hydrazoic acid. Certain organic azides, such as phenylazide, may be reduced in ethereal solution by means of stannous chloride and hydrogen chloride to the corresponding monosubstituted triazenes (88). Although triazene itself has never been isolated, indications of its existence in aqueous solution have been obtained. Dimroth and Pfister (89) treated a concentrated solution of potassium azide, containing some ammonium chloride and free ammonia, with zinc-copper powder at  $-10^{\circ}\text{C}$ . and found that immediate reduction took place. Filtration gave a solution which continued to evolve gas for some time in spite of the fact that the reducing agent had been removed. This behavior they assumed was due to the transitory existence of very unstable triazene.

Many reactions which might be expected to give triazene yield only ammonia and nitrogen. In water such reducing agents as stannous chloride-hydrogen chloride (89), hydriodic acid (40), sodium bisulfite (180b), and cuprous oxide (241) give these products. It may therefore be assumed that under the conditions employed in these cases reduction to triazene is followed by decomposition into nitrogen and ammonia.



A few trisubstituted triazenes have been reduced to the corresponding triazanes (234), a group of rather ill-defined substances and derivatives of the hydronitrogen,  $\text{N}_3\text{H}_5$ . From analogy with the corresponding reduction product of nitric acid, dihydroxylimine, which undergoes desolvation through the nitroxyl radical to give the dimer, hyponitrous acid, triazene might be expected to lose ammonia with formation of the  $[=\text{N}\cdot\text{NH}_2]$  radical or its tautomer, diimide. The reverse procedure, that of solvation, has been carried out in the preparation of certain organic derivatives of triazene, which are obtained by condensation of azo compounds with amines (87).

These considerations immediately call to mind the reduction of sodium azide in alkaline solution by metallic aluminum. Raschig (182b) found that one-third of the nitrogen is obtained in the form of ammonia and that variable quantities of hydrazine are obtained along with hydrogen and

nitrogen in approximately equivalent volumes. Raschig assumes that the decomposition of triazene yields nitrogen in an active form which is reduced to diimide, only to reappear again immediately in the form of free nitrogen and hydrogen. That diimide is actually an intermediate reduction product seems quite likely, especially since nitrogen and hydrogen have always been obtained in those cases where attempts have been made to prepare diimide.

However, it seems much more probable that reduction by aluminum yields principally triazane, which may then desolvate to give diimide, the latter decomposing into nitrogen and hydrogen. It should be pointed out that the  $[=N\cdot NH_2]$  radical might conceivably polymerize to give the dimer, tetrazene, which, as has been noted previously, decomposes to give hydrazine and nitrogen. The mechanism outlined here suggests the possibility of preparing these two hydronitrogens as such by the reduction of hydrazoic acid.

Ammonia and hydrazine are the common reduction products of hydrazoic acid. Metals (241, 244) react with hydrazoic acid to give nitrogen, ammonia, and variable quantities of hydrazine. Larger yields of hydrazine are obtained in those cases where reduction results in the separation of an insoluble hydrazine compound. Thus, reduction with zinc and sulfuric acid (56) results in the formation of an insoluble precipitate of the double salt,  $ZnSO_4 \cdot (N_2H_4)_2 \cdot H_2SO_4$ . Electrochemical reduction (241, 244) yields small quantities of these same substances. It is quite probable that the discrepancies in the theoretical ratios of nitrogen to hydrogen formed in the electrolysis of solutions of hydrazoic acid and azides, may be accounted for in this way. Sodium amalgam (56, 68), sodium polysulfide (56), and ferrous hydroxide (56) give mainly ammonia and little hydrazine. Both titanous and chromous chlorides (180) yield exactly 1 mole of ammonia per mole of hydrazoic acid used. Molecular and active hydrogen in the presence of colloidal palladium give ammonia and hydrazine. There is no appreciable reduction under these conditions in alkaline solution (191).

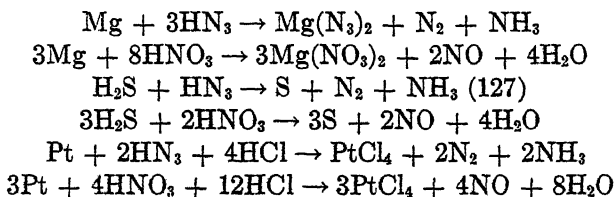
A critical survey of the chemistry of hydrazoic acid certainly indicates that the information concerning its behavior towards reducing agents, while prolific, is nevertheless fragmentary and in need of more extended study and investigation. Especially should the reduction of hydrogen azide in non-aqueous solvents be studied with the possible stabilization of such products as triazene and diimide in mind.

#### *G. Behavior as a nitridizing agent*

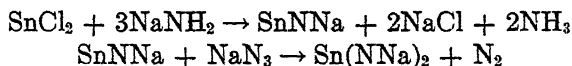
Attention has already been called to the fact that hydrazoic acid may be considered an *ammononitric* acid. It is formally derivable from ortho-

ammononitric acid by a process of deammonation (see p. 176). Potassium azide, an ammononitrate, is obtainable by the ammonolysis of potassium aquonitrate (100). Fused sodium amide reacts with sodium nitrate to give good yields of sodium azide (53). Lead azide can be obtained in 80 per cent yield by the action of an excess of potassium amide upon lead nitrate in liquid ammonia (100). All of these reactions may be considered to involve the conversion of an *aquonitrate*, a derivative of nitric acid, into an *ammononitrate*, a derivative of hydrazoic acid.

Just as nitric acid is characterized by its powerful oxidizing properties, so hydrazoic acid has been shown to be a powerful nitridizing agent. When effecting nitridation (a specific case of deelectronation), it undergoes reduction to give products which are analogous to the reduction products of nitric acid (see p. 190). Reaction with metals rarely gives hydrogen, but yields the corresponding metallic azides along with variable quantities of nitrogen, ammonia, and hydrazine (see p. 192). Turrentine (242) offers a series of parallel reactions to bring out the similarities between hydrazoic and nitric acids as deelectronators.

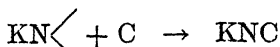
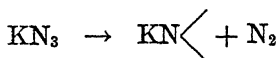


In a recently published article Franklin (100) summarizes various other experimental facts to support the assumption that hydrazoic acid is a powerful oxidizing agent. Ferrous azide is oxidized to ferric azide by warming with an excess of hydrazoic acid. Sulfur is converted to sulfuric acid by hot aqueous hydrazoic acid. Sodium ammonostannite, prepared by dissolving stannous chloride in fused sodium amide, is converted to the ammonostannate by action of sodium azide.

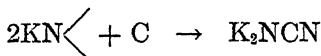


Fused potassium azide, like potassium nitrate, is a very efficient deelectronator. Carbon and potassium azide give potassium cyanide (101); in the presence of an excess of potassium azide the dipotassium salt of cyanamide (101) is obtained. This reaction is considered to involve the nitridation of an ammonocarbonite to an ammonocarbonate. In all of these reactions

the primary formation of the highly reactive  $\text{>NH}$  (or  $\text{>NK}$ ) group may be assumed. Thus



or



Nitridation of methylamine, an ammonioalcohol, to guanidine, an ammonocarbonic acid, is effected by heating methylamine with a liquid ammonia solution of ammonium azide. Numerous other cases involving nitridation of organic compounds are reported by Franklin (105) and by Schmidt (196).

#### *H. Detection and determination*

The most sensitive qualitative test for hydrazoic acid and azides in general is the blood-red coloration produced upon the addition of a solution of a ferric salt (72). This color, due to the formation of ferric azide, will last indefinitely, provided the ferric salt is present in excess. The test is sensitive to one part in 100,000 (80, 81) and has been made the basis of a colorimetric method for the quantitative determination of azides (145). The ferric azide color is distinguishable from that produced by ferric thiocyanate in two ways (80, 81). Dilute mineral acids cause the color of ferric azide to be discharged readily. The addition of mercuric chloride causes the color of ferric thiocyanate to fade very much more readily than that of the azide. Even insoluble metallic azides respond to the ferric azide test.

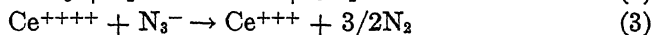
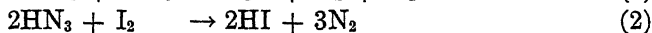
In the absence of other halogen or halogenoid ions, the addition of silver nitrate to a solution of a soluble azide causes precipitation of the extremely explosive silver azide.

A variety of methods has been used and proposed for the quantitative determination of azides. Curtius (72) used the absolute Dumas method for those azides which were not too explosive or too volatile. In their study of the reduction of hydrazoic acid Dennis and Isham (83) investigated the Kjeldahl method, but found that results were invariably low, probably because of loss of hydrogen azide upon contact of the sample with concentrated sulfuric acid. Since only one-third of the azide nitrogen is obtained as ammonia, this method is hardly to be recommended, even if loss of hydrogen azide could be avoided.

Curtius (72) analyzed soluble azides by precipitation as silver azide from

solutions made slightly acid with nitric acid. The precipitate was washed and then dried at 100°C. Results were usually low, owing to the fact, as Dennis (83) pointed out later, that silver azide is appreciably soluble in nitric acid. However, silver azide can be precipitated quantitatively from neutral solutions, or from solutions containing free acetic acid (83). Owing to the extremely treacherous nature of silver azide, Dennis and coworkers (78, 79, 82) converted the silver azide to silver nitrate by the action of nitric acid and then precipitated the silver in solution as silver chloride (31).

These difficult and hazardous methods have been displaced by simpler and more practical ones, involving either (1) the interaction of nitrous acid with hydrazoic acid, (2) the oxidation of azides by iodine in the presence of certain catalysts, or (3) oxidation by ceric salts.



In carrying out the analysis of azides according to equation 1, the sodium nitrite and the azide are allowed to react in a solution made acid with dilute sulfuric acid and the resulting gases are collected and measured (206, 207, 229, 250). Since nitrous oxide is appreciably soluble in water, results are apt to be low. A recently published method makes use of this same reaction, with the modification that the slightly acid solution of the azide is titrated with standard nitrite solution (187). Ferric chloride is used as the indicator and the end point is shown by a change from the deep red color of ferric azide to yellow.

One of the most useful gas-volumetric methods is based upon the oxidation of azides by iodine solutions in the presence of sulfides (185), thiosulfates (185), azidodithiocarbonates (39), or carbon disulfide (95). This method does not give accurate results when very dilute solutions are to be analyzed. However the procedure is very simple, involving only the addition of one of the above substances to a mixture of a potassium iodide-iodine solution and the azide to be analyzed under such conditions that the nitrogen gas formed may be collected and measured. If a standard iodine solution is used, the excess may also be determined by titration with standard sodium arsenite (119, 120).

Ceric salts also oxidize the azide nitrogen quantitatively to molecular nitrogen. The nitrogen gas may be collected and measured (57, 207), or the excess standard ceric salt solution may be determined, for instance, by the addition of potassium iodide and subsequent titration of the liberated iodine with standard thiosulfate using starch as indicator (155).

The acid character of hydrazoic acid has also been made the basis of

several analytical methods. Curtius (72) decomposed various inorganic azides with dilute sulfuric acid, collected the aqueous distillate of hydrazoic acid, and titrated the latter with a standard solution of potassium hydroxide using phenolphthalein as an indicator. West (252) varied this procedure by decomposing the salt with an excess of decinormal sulfuric acid, boiling off the hydrogen azide, and titrating the excess of sulfuric acid with standard alkali. Naturally, the hydrazoic acid distillate may be caught in standard alkali and the excess of the latter determined by titration with standard acid (57, 31).

#### VI. AZIDODITHIOCARBONIC ACID AND AZIDOCARBONDISULFIDE

Sommer (206) was the first to point out that, regardless of which structural formula is assigned to hydrazoic acid, the presence of the imide grouping should give to it properties characteristic of the secondary amines. Thus, hydrazoic acid and its salts (60) react with carbon disulfide to give azidodithiocarbonic acid and azidodithiocarbonates.



The acid and its salts are of particular interest since their oxidation leads to the formation of the free halogenoid (248) radical, azidocarbonbisulfide,  $(\text{SCSN}_3)_2$ , one of a group of inorganic radicals possessing halogen-like characteristics. Most of our knowledge concerning azidodithiocarbonic acid and its derivatives is the result of investigations by A. W. Browne and his coworkers.

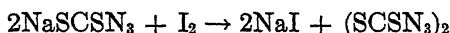
*Azidodithiocarbonic acid* may be prepared either by direct interaction of hydrazoic acid with carbon disulfide (170), or, better, by treatment of a concentrated solution of sodium azidodithiocarbonate with concentrated hydrochloric acid (205). In the latter case the free acid is obtained as a white crystalline solid, fairly soluble in water, but more soluble in the common organic solvents. It has been found to be a fairly strong acid in aqueous solutions and may be titrated (52) with standard alkali using methyl orange or methyl red as indicator. Measurement of the electrical conductance (203) of azidodithiocarbonic acid places the compound in acid strength above hydrofluoric, hydrazoic, and hydrocyanic acids, but below the halogen hydracids. Potentiometric titration shows that it is comparable in strength with sulfuric acid. The dry substance is very sensitive to both shock and heat and decomposes at room temperatures. In aqueous solution it is much more stable. In either case primary decomposition proceeds in accordance with the equation:



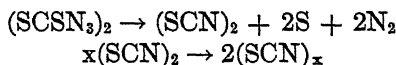
The thiocyanic acid formed rapidly undergoes polymerization and reaction to give products of indefinite composition and structure.

A whole series of salts (8, 36, 37, 41, 204) of azidodithiocarbonic acid has been prepared either (1) by direct interaction of the corresponding azides with carbon disulfide, (2) by the action of the acid upon the oxides and carbonates, or (3) by precipitation of the insoluble salts from aqueous solution. The silver, mercurous, and lead salts are insoluble in water (204). Not only are the heavy metal salts extremely explosive (186), but even such water-soluble salts as cesium azidodithiocarbonate are apt to decompose violently merely by contact during the process of crystallization from aqueous solution. Like the acid these salts undergo slow decomposition with formation of the corresponding thiocyanates, free nitrogen, and sulfur. The ammonium and substituted ammonium (8), as well as the potassium, rubidium, and cesium salts (36) are characterized by their peculiar sensitivity to light. All of these salts change color when exposed to sunlight. In some cases (cesium and rubidium salts) this color change is completely reversible. This phototropic effect is particularly noticeable in the case of the cesium salt, which turns a deep purple in strong sunlight. Investigation has shown that a rather limited region in the near ultra-violet (3750–3950 Å. U.) causes this change to take place.

*Azidocarbonedisulfide*,  $(\text{SCSN}_3)_2$  (39, 206), may be prepared by chemical or electrochemical oxidation of the free acid or one of its soluble salts. An iodine-potassium iodide solution is immediately decolorized when added to a solution of sodium azidodithiocarbonate with the resultant precipitation of the free dimeric radical as a white, microcrystalline, explosive solid (42).



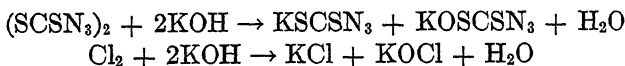
Azidocarbonedisulfide decomposes at room temperatures fairly rapidly, at higher temperatures with explosive violence, to yield as final products nitrogen and an intimate mixture of polymeric thiocyanogen and sulfur. Since this decomposition (38)



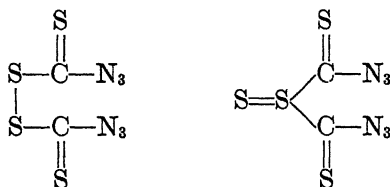
has been found to be autocatalytic, it is thought that the acceleration in the rate of decomposition of the halogenoid, as noted during the course of nitrometric studies, is due to the intermediate formation of an unstable interhalogenoid complex, possibly  $\text{SCN} \cdot \text{SCSN}_3$ . The halogenoid dissolves in potassium hydroxide to give a solution which upon acidification reprecipitates the original compound. On the basis of this and other chemical



evidence it has been assumed that the reaction occurring under these conditions is analogous to the well-known reaction which takes place when chlorine is dissolved in alkaline solution.



The oxyazidodithiocarbonate is not stable and such solutions gradually decompose with formation of a host of products, among which the thiocyanate, sulfate, thiosulfate, sulfide, and sulfite ions, in addition to free nitrogen and sulfur, have been identified. Azidocarbondisulfide in contact with water slowly undergoes simultaneous hydrolysis and decomposition to give similar products. The halogenoid is assumed to exist in two tautomeric forms, a disulfide and a monosulfide, to which the following structures have been assigned:



That azidocarbondisulfide does indeed possess halogenoid characteristics is further demonstrated by the fact that it forms the exceedingly unstable halogen derivatives,  $\text{ClSCSN}_3$ ,  $\text{BrSCSN}_3$ , and  $\text{Br}_3\text{SCSN}_3$  (107). The interhalogenoid, *cyanogen azidodithiocarbonate*,  $\text{CNSCSN}_3$  (7), has also been prepared either by the interaction of the potassium salt with ethereal cyanogen bromide, or by the action of the halogenoid upon mercuric cyanide. It is a white crystalline solid which melts with decomposition at  $67^\circ\text{C}$ . It undergoes decomposition to form polymerized cyanogen thiocyanate, nitrogen, and sulfur. When heated in inert solvents some cyanogen thiocyanate can be obtained.

Azidocarbondisulfide oxidizes hydrogen azide (256) in ethereal solution to give nitrogen. It undergoes simultaneous decomposition with formation of thiocyanogen, sulfur, and a variety of other products.

## VII. METALLIC AZIDES: AZIDES, TRINITRIDES, PERNITRIDES, AZOIMIDES, TRIAZOATES

### A. Preparation and general properties

Direct methods for the preparation of certain metallic azides have already been described in section II dealing with the synthesis of the azide radical. However, only two of these have been adapted to laboratory or

larger scale production; namely, the nitrous oxide-sodium amide method (80, 81, 257, 223, 138) and the alkyl nitrite-hydrazine synthesis (255, 166, 229, 215, 216). Both of these methods are used chiefly in the preparation of sodium azide, although they may be modified to obtain other azides. Thus, directions are given by Hoth and Pyl (138) for the preparation of potassium azide from potassium amide and nitrous oxide. These investigators also obtained 30 per cent yields of barium azide from ethyl nitrite, hydrazine hydrate, and barium hydroxide. In most instances, however, sodium azide must serve as the starting material for the preparation of all other salts.

Solutions of sodium azide react metathetically with solutions of soluble copper (72), silver (61), lead (62), mercurous (61, 62), and thallous (82) salts to precipitate the corresponding azides. All other common metallic azides are soluble in water and are obtainable (1) by the action of hydrazoic acid upon the metals, oxides, hydroxides, or carbonates, (2) by metathesis of barium azide with soluble sulfates, and (3) by double decomposition of potassium azide with a soluble perchlorate. Since silver azide is much more soluble than silver iodide, Friedlander (104) was able to use the reaction between tetramethylammonium iodide and silver azide to obtain tetramethylammonium azide. The extremely treacherous character of silver azide militates seriously against the general adoption of this method.

In many cases metallic salts are not obtainable from aqueous solution as such, but give on standing or evaporation insoluble basic salts, some of extremely variable composition. Such azides as those of magnesium (72), zinc (72), yttrium (68), lanthanum (68), cerous cerium (68), chromium (68), manganese (68, 72), and cobalt and nickel (72) yield basic compounds upon evaporation of aqueous solutions of the corresponding salts. In the case of aluminum (72), zirconium (68), and thorium<sup>5</sup> (68) precipitation of the corresponding hydroxides takes place even in the cold when sodium azide is added to solutions of salts of these metals in water. Quantitative

<sup>5</sup> Dennis and Kortright (85, 78) made use of potassium azide for the separation of thorium from the rare earths of the cerium and yttrium group, and recommend this reagent as one which may be used both for the qualitative detection of thorium and for its quantitative determination either alone or in the presence of other rare earths. Recently, this method has been extended to the separation of the rare earths themselves (3). Cerium can be oxidized to the ceric state readily by hydrogen peroxide. Addition of sodium azide is said to result in the quantitative separation of ceric hydroxide, leaving the other rare earths in solution. If such a solution is boiled the less basic rare earths precipitate first, with the result that the residual solution is gradually enriched with respect to the more basic members. J. Ant-Vuorinen (3) claims to have prepared pure lanthanum material in this manner. This same investigator has also shown that samarium hydroxide is precipitated quantitatively when solutions containing samarium salts and sodium azide in excess are boiled (144).

precipitation of ferric hydroxide (68) is effected when a solution of the azide is boiled. Hahn (121) recommends the use of sodium azide-sodium nitrite solutions for the precipitation of the hydroxides of iron, aluminum, and chromium in a granular and easily filterable form.

Non-aqueous media have been used successfully in the preparation of certain azides. Thus, ammonium (106), ethylammonium, and diethylammonium (129) azides may be obtained by interaction of ethereal solutions of hydrogen azide with the corresponding anhydrobases in ether or alcohol. Dihydroxylammonium azide,  $(\text{NH}_2\text{OH})_2 \cdot \text{HN}_3$  (83), is best prepared in alcoholic solution from its components. Zinc azide,  $\text{Zn}(\text{N}_3)_2$  (263), is obtained by action of ethereal hydrogen azide upon metallic zinc. Chromium azide,  $\text{Cr}(\text{N}_3)_3$  (173), may be prepared by the reaction of anhydrous chromium nitrate with sodium azide in anhydrous alcohol. Ferric azide,  $\text{Fe}(\text{N}_3)_3$  (263), is obtainable when dry ferric sulfate and sodium azide are shaken together in absolute methanol. The precipitated sodium sulfate is removed by filtration and the solution of ferric azide concentrated in a vacuum desiccator. Reaction of cobaltous carbonate with ethereal hydrogen azide gives the anhydrous cobaltous azide,  $\text{Co}(\text{N}_3)_2$  (263).

The alkali and alkaline earth metals react with ammonium azide in liquid ammonia to give the corresponding azides (45, 98). Browne and his coworkers (44) have prepared various metallic azides by electrolyzing solutions of ammonium azide in liquid ammonia, using the respective metals as anodes. No analyses are recorded for the products obtained, but they claim to have prepared the normal azides of silver, cadmium, lead, and antimony by this method. With aluminum, iron, and nickel anodes ammono-basic azides are presumably obtained.

Stabilization of some metallic azides is effected by converting them into ammonates, pyridinates, and even hydrazinates (90). Thus, zinc azide dipyridinate,  $\text{Zn}(\text{N}_3)_2 \cdot 2\text{C}_5\text{H}_5\text{N}$  (83), which is precipitated upon addition of pyridine to an aqueous solution of the azide, is much more stable than zinc azide itself. Pyridinates of cadmium, cobaltous, and nickelous azides have been prepared in a similar manner.

Double salts, such as  $\text{Ni}(\text{N}_3)_2 \cdot \text{NH}_4\text{N}_3$  and  $\text{Ni}(\text{N}_3)_2 \cdot \text{KN}_3$  (72, 263), are obtained by adding the calculated quantities of the ammonium or potassium azides to concentrated aqueous solutions of nickelous azide. The corresponding cobaltous salts (72) have also been prepared.

It is rather questionable whether the compound  $\text{Cr}(\text{N}_3)_3 \cdot 3\text{NaN}_3$  (173) is to be considered a double salt. It is probably the sodium salt of a complex chromihydrazoic acid,  $\text{H}_3\text{Cr}(\text{N}_3)_6$ , since it yields with silver nitrate solution not silver azide, but a precipitate whose composition has not been determined because of its highly explosive nature. It is significant in this

connection that Mendelejeff (158) predicted the existence of complexes of this type some forty years ago, shortly after the discovery of hydrazoic acid by Curtius.

Only recently, however, has an attempt been made to prepare and describe cobalt complexes containing the azido grouping, although reference to such an investigation may be found in Mellor.<sup>6</sup> Strecker and Oxenius (218) were unable to apply the usual methods for the preparation of such complexes, because of the tendency for cobalt azide to hydrolyze. However, interaction of hexamminocobaltic sulfate with barium azide in aqueous solution yields hexamminocobaltic azide,  $[\text{Co}(\text{NH}_3)_6](\text{N}_3)_2$ . From chloropentamminocobaltic sulfate and barium azide is obtained chloropentamminocobaltic azide,  $[\text{Co}(\text{NH}_3)_5\text{Cl}](\text{N}_3)_2$ . These compounds are both soluble in water. Treatment of tetramminocarbonatocobaltic chloride with hydrazoic acid gives tetramminodiazidocobaltic azide,  $[\text{Co}(\text{NH}_3)_4(\text{N}_3)_2]\text{N}_3$ . The corresponding compound, diethylenediamminodiazidocobaltic azide,  $[\text{Co}(\text{en})_2(\text{N}_3)_2]\text{N}_3$ , is prepared in a similar manner. Several other compounds of this type are described by Strecker and Oxenius.

Ricca and Pirrone (192) have prepared a series of addition compounds of azides with mercuric cyanide having the general formula  $\text{Hg}(\text{CN})_2 \cdot \text{Me}(\text{N}_3)_2$  (where  $\text{Me} = \text{Cu}^{++}, \text{Zn}^{++}, \text{Co}^{++}, \text{Ni}^{++}$ ). Vournasos has also prepared several rather remarkable compounds of the complex type by allowing sodium azide to react with antimony tribromide in acetone to give  $\text{NaSb}_3\text{Br}_9\text{N}_3$  (247) and with arsenic tribromide in methyl alcohol to give  $\text{Na}_3(\text{AsBr}_3\text{N}_3)_3$  (246).

### B. Properties

While the chemical properties of the metallic azides are fairly well known, since they resemble in most respects those characteristics of the azide radical discussed under hydrazoic acid, little is known of the physical properties of these compounds. Only the alkali and alkaline earth azides have been the subjects of more than a cursory examination. The physical properties of these azides have been collected in table 1. The physical properties of those heavy metal azides which have found application as detonators are included under the discussion of the explosive properties of azides (see p. 207).

<sup>6</sup> Mellor: A Comprehensive Treatise on Inorganic and Theoretical Chemistry, Volume VIII, p. 355. The preparation of a series of cobalt complexes by S. Lorie is noted from information obtained from a publication by this author entitled "Stickstoffwasserstoffsäure und ihre anorganischen Verbindungen," Zürich (1912).

TABLE 1  
Physical properties of some inorganic azides

FORMULA	MELTING POINT	DECOMPOSITION TEMPERATURE	DECOMPOSITION TEMPERATURE* (235)	SOLUBILITY IN 100 g. WATER†	SOLUBILITY‡ IN OTHER SOLVENTS PER 100 g. SOLVENT	CRYSTAL FORM	OTHER PHYSICAL PROPERTIES	SOLVATES
	degrees C.	degrees C.	degrees C.	grams	grams			
LiN <sub>3</sub> .....				66.41 <sup>145</sup>	C <sub>2</sub> H <sub>5</sub> OH: 20.26 <sup>145</sup> Insoluble in ether			LiN <sub>3</sub> ·H <sub>2</sub> O (79) LiN <sub>3</sub> ·2H <sub>2</sub> O (270) LiN <sub>3</sub> ·4H <sub>2</sub> O (270)
NaN <sub>3</sub> .....		330 (235) 275 (219) 300 (162)	280	41.71 <sup>75</sup>	C <sub>2</sub> H <sub>5</sub> OH: 0.22 <sup>95</sup> (58) Soluble in liquid NH <sub>3</sub> (105)	Hexagonal (72, 130) <sup>†</sup>	d <sub>4</sub> <sup>25</sup> = 1.8473 (161, 58, 130) Refractive index, aqueous solution (58) Conductivity, aqueous solution (58) Density, aqueous solution (31) Viscosity, aqueous solution (31)	
KN <sub>3</sub> .....	343 (219) 350 (135)	355 (219) 320 (235)	360	49.61 <sup>75</sup>	C <sub>2</sub> H <sub>5</sub> OH: 0.14 <sup>145</sup> Soluble in liquid NH <sub>3</sub> (105)	Tetragonal (79, 130) <sup>†</sup>	d = 2.056 (58, 130) Refractive index, aqueous solution (58) Conductivity, aqueous solution (58)	
RbN <sub>3</sub> .....	300 (235) 321 (219) 330 (72)	260 (235) 385 (219)	310	114.11 <sup>75</sup>	Slightly soluble in C <sub>2</sub> H <sub>5</sub> OH Insoluble in ether	Tetragonal (79, 176, † 54) <sup>†</sup>	d = 2.9885 (118)	
CsN <sub>3</sub> .....	320 (235) 326 (219)	290 (235) 390 (219)	350	307.41 <sup>45</sup>	C <sub>2</sub> H <sub>5</sub> OH: 1.04 <sup>145</sup> Insoluble in ether	Tetragonal (79)		
NH <sub>4</sub> N <sub>3</sub> .....	Sublimes (106, 72)			13.3 <sup>95</sup> 20.16 <sup>95</sup> 27.07 <sup>95</sup> (106)	CH <sub>3</sub> OH: 3.27 <sup>95</sup> C <sub>2</sub> H <sub>5</sub> OH: 1.06 <sup>95</sup> ; 1.39 <sup>95</sup> Insoluble in ether, C <sub>6</sub> H <sub>6</sub> Very soluble in liquid NH <sub>3</sub> (106)	Rhombic (72)	d <sub>4</sub> <sup>25</sup> = 1.3459 (106) Vapor pressure of solid (106) Density of saturated solutions (106) Thermal data (18, 19)	NH <sub>4</sub> N <sub>3</sub> ·NH <sub>3</sub> NH <sub>4</sub> N <sub>3</sub> ·5NH <sub>3</sub> (91)
N(CH <sub>3</sub> ) <sub>3</sub> N <sub>3</sub> (104).....		125		Very soluble	Soluble CH <sub>3</sub> OH, C <sub>2</sub> H <sub>5</sub> OH Slightly soluble in C <sub>6</sub> H <sub>6</sub> , CHCl <sub>3</sub> , ether	Tetragonal		
(NH <sub>4</sub> OH) <sub>2</sub> HN <sub>3</sub> (83).....	66			Very soluble	Soluble in C <sub>2</sub> H <sub>5</sub> OH Insoluble in ether			
N <sub>3</sub> H <sub>4</sub> ·HN <sub>3</sub> .....	75.4 (91)			Very soluble	N <sub>2</sub> H <sub>4</sub> : 190 <sup>95</sup> CH <sub>3</sub> OH: 6.1 <sup>95</sup> C <sub>2</sub> H <sub>5</sub> OH: 1.6 <sup>95</sup> (91)	Rhombic (91)		N <sub>3</sub> H <sub>4</sub> N <sub>3</sub> ·N <sub>2</sub> H <sub>4</sub> (81, 193, 53)
Ca(N <sub>3</sub> ) <sub>2</sub> .....		110 (235)	100	45 <sup>53, 55</sup>	C <sub>2</sub> H <sub>5</sub> OH: 0.211	Rhombic (79)		Ca(N <sub>3</sub> ) <sub>2</sub> ·2N <sub>2</sub> H <sub>4</sub> Ca(N <sub>3</sub> ) <sub>2</sub> ·N <sub>2</sub> H <sub>4</sub> (90)
Sr(N <sub>3</sub> ) <sub>2</sub> .....		140 (235)	110	45.83 <sup>145</sup>	Slightly soluble in C <sub>2</sub> H <sub>5</sub> OH	Rhombic (79)		
Ba(N <sub>3</sub> ) <sub>2</sub> .....		160 (235)	120	16.73 <sup>95</sup>	Slightly soluble in C <sub>2</sub> H <sub>5</sub> OH	Monoclinic (118) <sup>†</sup>		Ba(N <sub>3</sub> ) <sub>2</sub> ·H <sub>2</sub> O (80, 81)

\* The figures given in column 3 refer to those temperatures at which regular nitrogen evolution takes place after decomposition has begun (see p. 205).

† Unless otherwise noted all solubilities were determined by Curtius and Rissom (72).

‡ Indicates an x-ray crystal structure study.

§ Solubilities of ammonium azide, both in water and in other solvents, refer to grams per 100 grams of solution

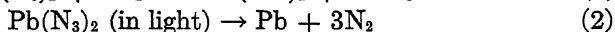
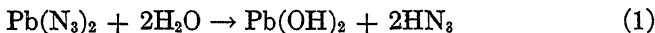


### *C. Photochemical decomposition of azides*

In his first description of silver azide, Curtius (61) stated that it was possible to distinguish this compound from the halides of silver by the fact that it was not light-sensitive. Subsequent investigations by Wöhler and Krupko (260, 261) and by Bekk (13, 14) showed that this was not the case. Not only silver azide but several other metallic azides undergo distinct decomposition upon exposure to light. Silver azide darkens slowly with formation of metallic silver and evolution of nitrogen, a process which may be followed both microscopically where particles of silver may be detected readily, and also gasometrically, by measuring the volume of nitrogen gas evolved. Bekk (13, 14) prepared dry emulsions of the silver salt and found that they were insensitive to shock and could be manipulated without danger. It is claimed that such an emulsion is much more easily reduced by developing solutions than the corresponding halides, and that it is characterized by its relatively high sensitivity to the red portion of the spectrum. More recently (4), however, it has been found that emulsions of silver azide absorb light in the same regions as silver chloride. This absorption does not parallel that of the chloride, but decreases in the region of shorter wave length.

Mercurous azide is extremely light-sensitive and takes on a superficial yellow coloration within a few minutes. The formation of droplets of free mercury is readily discernible under the microscope. Cuprous azide is also decomposed by sunlight.

Dry lead azide assumes a grayish yellow color on exposure to light. Under water the color changes are much more marked, and decomposition takes place with formation of appreciable quantities of ammonia. This reaction does not take place in the dark, so that the reduction of hydrazoic acid liberated by hydrolysis of the lead salt may be said to be due to the action of the metallic lead formed by the photochemical decomposition of the azide. The following reactions probably take place.



Sodium azide is decomposed by ultra-violet light at wave-lengths below 4050 Å. U. (163). When barium azide is exposed to soft x-rays decomposition takes place with formation of barium nitride (116).<sup>7</sup>

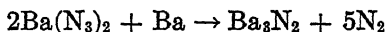
<sup>7</sup> Müller and Brous (163, 164) have studied the decomposition of sodium azide by bombardment with a stream of electrons. Decomposition with evolution of nitrogen does not take place until the electrons have assumed an energy of  $12 \pm 0.1$  volts. Muraour and Trillat (269) observed darkening of lead azide upon bombardment with electrons.

*D. Thermal decomposition of azides*

All azides are characterized by their relative instability towards heat. At higher temperatures all of them decompose with more or less violence to give free nitrogen and the corresponding metal, sometimes with varying quantities of the nitride. Very few azides are stable at their melting points, and even these undergo decomposition if kept in the fused state over longer periods of time. In table 1 are given the melting points and decomposition temperatures of the alkali and alkaline earth azides. In the third column are noted the temperatures at which, according to Tiede (235), regular evolution of nitrogen takes place once decomposition has begun. The temperatures given by Tiede and by Suhrmann and Clusius (219) refer to observations made in a vacuum. Determinations of the melting points and decomposition temperatures made by other authors were carried out at atmospheric pressures.

It is quite evident that the temperature required to initiate decomposition of certain azides is higher than that required to carry it forward, once it has begun. Tiede ascribed this behavior to the catalytic effect of the liberated metals. This assumption has since been verified by Günther, Andreew, and Ringbom (115), who were able to prove that the explosive decomposition of barium azide is due, not necessarily to direct dissociation of the azide into metal and nitrogen, but rather to the secondary reaction of metallic barium with barium azide to give the nitride and free nitrogen.

Garner and Moon (109), as well as Harvey (128), have shown that the small particles of metallic barium first formed on heating barium azide act as nuclei and catalyze the further decomposition in accordance with the equation:



Curiously enough, exposure of crystals of barium azide to radium emanation shortens very appreciably the time required to bring about initial decomposition. An induction period (109) is necessary to start decomposition even under ordinary conditions, but according to Garner and Moon (109) the time is reduced at 110°C. from 110 to 55 minutes and the speed of decomposition tripled by exposure to radium emanation of intensity of 1 millicurie.

The heavy metal azides tend to undergo sudden and explosive dissociation over a range of temperatures depending upon crystal size, purity, rate of heating, and the presence or absence of admixed substances. Their explosive properties will be discussed later. Hitch (132) carried out a series of experiments effecting slow decomposition by using extremely small quantities (0.05 g.) of various metallic azides. The initial decomposition



temperatures and explosion temperatures as determined by Hitch are given in table 2.

Tiede (235, 237) recommends the thermal decomposition of the alkali and alkaline earth azides as a method for the preparation of very pure nitrogen. Justi (140) also decomposed sodium azide to prepare very pure nitrogen. Booth (30) determined the atomic weight of nitrogen by decomposing silver azide. An average of fourteen determinations of the silver-nitrogen ratio gave an atomic weight of 14.007 for nitrogen. Hitch (132) determined the atomic weight of silver from the silver azide-silver ratio obtaining values varying between 107.93 and 107.98. Moles (162) used sodium azide for the determination of the atomic weight of sodium and obtained values whose extremes were 23.024 and 23.10. The fact that both Moles and Hitch obtained values higher than those generally accepted indicates that both determinations were subject to a common error.

TABLE 2  
*Decomposition temperatures of some heavy metal azides*

AZIDE	INITIAL DECOMPOSITION TEMPERATURE	EXPLOSION TEMPERATURE
	<i>degrees C.</i>	<i>degrees C.</i>
HgN <sub>2</sub> .....	215	270
Hg(N <sub>3</sub> ) <sub>2</sub> .....	215	300
Pb(N <sub>3</sub> ) <sub>2</sub> .....	250	360
TlN <sub>3</sub> .....	Melts at 330; sublimes at 340	430
Ba(N <sub>3</sub> ) <sub>2</sub> .....	180	225

The thermal decomposition of azides (92, 188, 219, 202, 27) has been used for the preparation of various metals in a state of high purity. Ebler (92) prepared a barium azide containing approximately 2 per cent radium azide and decomposed this preparation, obtaining a metallic mirror which had lost but little of its activity. Suhrmann and Clusius (219) studied the decomposition of the alkali azides as a means for preparing the alkali metals in very pure form. The azides were allowed to decompose slowly in a high vacuum over periods of three to four days at temperatures noted in table 1. Under these conditions the liberated metals distilled away from the residual masses in the decomposition vessel giving the following yields of pure alkali metal:

Azide.....	NaN <sub>3</sub>	KN <sub>3</sub>	RbN <sub>3</sub>	CsN <sub>3</sub>
Per cent yield.....	100	80	60	90

The residues remaining in the case of potassium, rubidium, and cesium yielded considerable quantities of ammonia upon treatment with water,

indicating that nitrides had been formed under these conditions. These residues were subsequently analyzed by Clusius (55), who verified this supposition. Since the stability of the azides rises with increasing atomic weight of the alkali metal, Suhrman and Clusius assumed that the nitrides probably exhibited a reverse order of stability, thus accounting for the relatively large amount of  $\text{Rb}_3\text{N}$  formed under the temperature conditions employed. Wattenberg (249) decomposed sodium azide partially and observed limited nitride formation under these circumstances. Tiede observed high percentage of nitride formation in the decomposition of the alkaline earth azides. Günther and his coworkers (115) found that yields of nitride as high as 75 per cent could be obtained by careful decomposition of barium azide.

De Boer, Clausing, and Zecher (77) decomposed mixtures of cesium chloride and barium azide thermally in a vacuum, and were able to obtain mirrors of metallic cesium by this method.

If the alkali and alkaline earth azides are decomposed in an atmosphere of hydrogen the corresponding hydrides may be prepared. Wattenberg (249) obtained sodium hydride in this manner, and Tiede (236) applied this method to the preparation of barium hydride.

#### *E. Explosive properties of azides*

While it is not the purpose of this article to discuss in detail the use of certain azides as initiators for the detonation of secondary explosives, it is nevertheless worth while to point out some of the interesting facts upon which these applications rest. The great sensitivity of various metallic azides to heat, impact, and friction suggested their possible use as detonants soon after the discovery of hydrazoic acid and the preparation of its inorganic derivatives by Curtius (62). As early as 1893 (154b, 258) the Prussian government investigated the azides of lead, silver, and mercury from this point of view, but a fatal accident resulted in the discontinuance of these experiments. Shortly thereafter, Berthelot and Vieille (19) directed attention to the explosive properties of mercurous and ammonium azides, pointing out that the latter compared favorably with guncotton insofar as temperature ( $1400^\circ\text{C}.$ ) attained during explosion and the pressure developed was concerned. One kilogram of ammonium azide upon decomposition yields 1148 liters of gas measured at standard conditions. However, nothing was done concerning the possible utilization of azides until Wöhler called attention to them about 1907 as possible substitutes for mercuric fulminate. After considerable investigation by Wöhler (258, 262), the manufacture of lead azide was begun in Germany about 1914 (189). It was not until the Italian advance upon Gorizia brought the Austrian mercury mines within range of the Allied guns that the Central

powers were forced to consider seriously the use of lead azide detonants (190). It is therefore fairly certain that lead azide was used during the World War by the Germans, who seem also to have been the only ones to have considered it seriously as a substitute for mercuric fulminate. Since 1920 the use of lead azide has developed considerably (253, 213, 225, 122). A large and extensive patent and technical literature is evidence of this growing interest in and application of azide detonators (274 to 279).

While mercuric fulminate is still the most important constituent of primers<sup>8</sup> it is not satisfactory in all respects. It deteriorates on long storage, especially if kept under unusual temperature conditions. It is somewhat hygroscopic and its effectiveness is very much decreased by the presence of moisture. It does not function satisfactorily when highly compressed. Even though lead azide has received most attention as a substitute for mercuric fulminate, it is superseded in initiatory power by

TABLE 3  
*Effectiveness of various azides and fulminates*

METAL	AZIDE	FULMINATE
Cadmium.....	1	2
Silver.....	3	4
Lead.....	5	—
Cuprous.....	6	7
Mercurous.....	8	—
Mercuric.....	—	9

several others. Marshall (154b) gives numbers to show in descending order the effectiveness of various azides and fulminates (see table 3).

Both Wöhler (211) and Birckenbach (23) have emphasized the extremely sensitive character of cadmium azide. It can not be employed as a detonant on account of its solubility in water and its hygroscopicity. The cost and extreme explosiveness of silver azide have acted as serious deterrents to its extended use. However, Taylor and Rinkenbach (226) have found that precipitation of silver azide from concentrated solutions of sodium

<sup>8</sup> In a private communication to the author L. V. Clark states, "The term 'primer' according to accepted nomenclature designates those devices used to effect the ignition of the propellant charge in small arms or large caliber military ammunition, or, in the case of high explosive shell, is the device which, acting through a booster charge of tetryl, effects the detonation of the main charge of explosive within the shell. It is true that mercuric fulminate is still the most important constituent of primers for the larger caliber military ammunition. However, in the case of primers for small arms ammunition, there has been a definite effort on the part of manufacturers to eliminate mercuric fulminate from their compositions."

azide and silver nitrate produces a semicolloidal material which is much safer to handle, much less hygroscopic, and more stable toward temperature changes, than the more crystalline material obtained from dilute solutions (see also 29). Silver azide is much more sensitive than either mercuric or silver fulminate and develops a higher initial impulse. This is true of several of the more sensitive azides, thus making it possible to use a much smaller quantity of azide than fulminate to detonate a given quantity of high explosive (see table 4).

Although lead azide (59, 141) (1) withstands unusual temperature conditions in storage, (2) is not appreciably affected by moisture, (3) develops a higher initial impulse than mercuric fulminate, and (4) does not become "dead pressed," its commercial development was retarded by the fact that it is slowly attacked by wet carbon dioxide and hydrogen sulfide (273) with the resultant liberation of hydrazoic acid, which in turn attacks

TABLE 4  
*Minimum detonating charge (154c)*

INITIATOR	MINIMUM CHARGES IN GRAMS WHICH WILL DETONATE 0.5 G. OF			
	Tetryl	Picric acid	T.N.T	Trinitroanisole
Hg(ONC) <sub>2</sub> .....	0.29	0.30	0.36	0.37
HgN <sub>3</sub> .....	0.05	0.08	0.15	0.55
Pb(N <sub>3</sub> ) <sub>2</sub> .....	0.03	0.03	0.09	0.28
AgN <sub>3</sub> .....	0.02	0.04	0.07	0.26
Cd(N <sub>3</sub> ) <sub>2</sub> .....	0.01	0.02	0.04	0.10

copper or brass detonator capsules to form the extremely treacherous cuprous azide. It has been found that if sufficient care is taken to dry thoroughly the ingredients entering into a detonator composition, the danger of decomposition of lead azide is eliminated.<sup>9</sup> The difficulty has also been overcome by use of aluminum or aluminum alloy materials as containers (276). Lead azide is also more difficult to ignite, although it is claimed that careful purification of the sodium azide and the water used in its preparation yields a better product. This difficulty has largely been overcome either by the use of composite azide detonators, where other ingredients are added to lower the ignition temperature, or an ignition mixture designed to develop a high temperature is superimposed upon the azide charge. Practically all patent specifications now call for the admixture of other substances in the manufacture of azide detonators (275). Efforts to use basic lead azide (258) or double salts of lead azide (26) have not been particularly successful.

<sup>9</sup> Private communication from Dr. L. V. Clark.

Only recently has it been shown that lead azide exists in two crystalline modifications (108, 159). The ordinary form (denoted as the  $\alpha$ -form) crystallizes in the orthorhombic system ( $d = 4.71$ ). It is much more stable and less sensitive to shock and friction than the  $\beta$ -form ( $d = 4.93$ ), which is monoclinic. This observation has been checked by determining the rate of decomposition of both forms from  $210^{\circ}\text{C.}$  to the explosion temperature ( $295^{\circ}\text{C.}$ ). The unstable  $\beta$ -form has a much lower critical increment than the  $\alpha$ -form (47,600 calories). Furthermore, it has been found that the  $\beta$ -form often explodes spontaneously, which may be the cause for some of the accidents which have occurred in the manufacture of lead azide. The  $\beta$ -form undergoes transition into the  $\alpha$ -form by contact with a solution of a lead salt.

Lead azide, like lead chloride, is appreciably soluble in warm water. Upon cooling lead azide sometimes comes out in the form of large crystals (probably the  $\beta$ -form (159)), which are extremely sensitive.<sup>10</sup> This fact is in accord with previous observation that, owing to internal strains, large crystals of various fulminates and azides are inherently much less stable than the finely divided product. Present technical methods and patent specifications call for the precipitation of lead azide in finely divided or colloidal form (151, 131, 214).

Mercurous azide (114) has also been proposed as a constituent of composite detonators. It is claimed that mercurous azide is not affected by carbon dioxide and may be used in copper detonator tubes. Mercuric azide, like lead azide, appears to exist in two crystalline modifications (159). The  $\beta$ -form is even less stable than the  $\beta$ -lead azide. Its solubility in water will probably prevent its consideration for use as a priming agent.

In addition to the azides of silver and lead and mercury, cyanuric triazide,  $(\text{CN})_3(\text{N}_3)_3$ , (175, 227, 141, 228, 259) has been mentioned prominently as a possible detonating agent. According to Kast and Haid (141) it develops a maximum detonation velocity of approximately 7500 meters per second when compressed to a density of 1.54, whereas the figure for lead azide with density of 4.6 is 5300 meters per second. Taylor and Rinkenbach (228) have studied the sensitiveness of various detonating compounds to frictional impact, impact, and heat. A glance at their figures (see table 5) will indicate immediately that cyanuric triazide possesses interesting possibilities. It is, however, relatively expensive to manufacture and rather volatile.

The sand-crushing strength test (see table 6) shows very clearly that

<sup>10</sup> The mechanism of and the various phenomena accompanying the explosion of lead azide under varying conditions have been the subjects of considerable study by Muraour and his coworkers (267, 268) as well as others (265, 271, 272).

cyanuric azide is probably the most brisant of the proposed and commercial initiators.

A comparative study by Muraour (165) of lead and cyanuric azides and mercuric fulminate is reproduced in table 7, in which the distinctive physical properties of these substances are listed.

TABLE 5  
*Sensitiveness to friction (228)*

AZIDE	PENDULUM FRICTION			IMPACT FALL 500 G. WEIGHT	EXPLOSION TEMPERA- TURE
	Added weight in kilograms	Fall	Number of swings		
					<i>degrees C.</i>
Pb(N <sub>3</sub> ) <sub>2</sub> .....	0.45	37.5	12	43	383
AgN <sub>3</sub> .....	4.35	33.0	30	41	273
HgN <sub>3</sub> .....	1.00	50.0	16	6	298
(CN) <sub>3</sub> (N <sub>3</sub> ) <sub>3</sub> .....	0.00	12.5	3	7	252

TABLE 6  
*Sand-crushing strength of certain of the azides*

WEIGHT OF CHARGE OF EXPLOSIVE	WEIGHT OF SAND CRUSHED FINER THAN 30 MESH BY			
	Mercury fulminate	Lead azide	Silver azide	Cyanuric triazide
<i>grams</i>	<i>grams</i>	<i>grams</i>	<i>grams</i>	<i>grams</i>
0.10	—	—	3.3	4.8
0.20	3.8	5.9	6.8	12.2
0.30	8.0	—	10.4	—
0.40	12.2	12.7	—	33.2
0.50	16.0	16.1	18.9	—
0.60	20.1	20.9	—	54.4
0.75	—	—	30.0	—
0.80	28.2	28.5	—	68.9
1.00	36.8	33.6	41.1	78.6

The use of sodium azide as a constituent of a priming charge has been patented (279). A recent study of the explosive properties of the alkaline earth azides has shown that calcium azide is the most sensitive (266).

An interesting study dealing with the sensitivity of various salts of hydrazoic acid to heat has been carried out by Wöhler and Martin (273). The data show very clearly the dependence of the true point of deflagration of the azides upon the quantity of substance. By plotting the minimum temperatures at which no explosion took place in 5 minutes against the

TABLE 7  
*Characteristics of certain explosives (165)*

EXPLOSIVE	H.F.M.  calories per molecule	H.P.G.  calories per gram	$T_d$  degrees C.	CHARGE LIMITS  grams	$Q$  kg.-cal. Hg vapor	$Q_t$  kg.- cal.	$V_0$  liters Hg vapor	VOL.  liters	$T_A$	$P$	$P_t$	CO- VOL- UME  cm. <sup>3</sup>	$V_E$  cm. <sup>3</sup>	RATE OF DETONA- TION  Meters per second	
														Density	
Mercury fulminate.....	-64.3	-223	169-175		368 Hg vapor	1360	315	4720	5630			315			
d = 4.43 (density of crystals).....						1395				24941			226	1.66	2750
d = 3.55 (350 atmospheres per cm. <sup>2</sup> )...						1306		1118		19986			282	3.30	4500
d = 3.30 (200 atmospheres per cm. <sup>2</sup> )...				0.80 (A) 0.20 (B)		1214		1039		18579			303		
Lead azide.....	-110.8	-380	305-320		260 Pb vapor		308	3720	4339			308			
d = 4.79 (density of crystals).....					1245			1475		20784			209	2.80	4500
d = 3.50 (800 atmospheres per cm. <sup>2</sup> )...				0.10 (A)	1078			970		15186			286	3.80	4500
d = 2.90 (200 atmospheres per cm. <sup>2</sup> )...				0.19 (A)	893			893		12583			345		
Cyanuric triazide.....	-219	-1073	197-206		1140* 700†		659*†	4220* 3320†	10590* 8286†			659			
d = 1.4 (200 atmospheres per cm. <sup>2</sup> ).....				0.12 (A) 0.08 (B)	1596* 980†			923		14742* 11600†			714	1.15	6550
d = 1.5 (800 atmospheres).....				0.20 (A)	1710* 1050†			988		15795* 12429†			667		

H.F.M. = heat of formation from the elements in large calories per molecule.

H.F.G. = heat of formation from the elements in small calories per gram.

$T_d$  = temperature of decomposition—the temperature at which the explosive decomposes on an oil bath.

Charge limit = minimum weight of the explosive which is able to detonate T.N.T. compressed to a density of 1.35 in a copper tube of the type used for the preparation of No. 8 detonators, the initiating explosive having been compressed on top of the T.N.T. The letter A signifies that the upper surface of the initiating explosive is free; the letter B signifies that on this upper surface has been placed a metallic capsule.

$Q$  = the number of large calories given off by 1 kg. of the explosive.

$Q_i$  = the number of large calories given off by the decomposition of 1 liter of the explosive.

$V_0$  = gaseous volume measured at 0° and 760 mm. (water vapor) given off by 1 kg. of the explosive.

Vol. = the volume of gas given off by the decomposition of 1 liter of the explosive.

$T_A$  = the absolute temperature attained by the gas on explosion.

$P$  = force of the explosive. The pressure which is developed by the gas given off by the decomposition of 1 kg. of the explosive in a volume of 1 liter if this gas follows the laws of a perfect gas at high pressure.

$P_i$  = the pressure which would be generated by the gas liberated by the decomposition of 1 liter of explosive detonated in its own volume if the gas again follows the laws of perfect gas.

Covolume of the gas of the explosion. Covolume has been given, according to Sarrau, to be 1/1000 of  $V_0$ .

$V_E$  = volume occupied by 1 kg. of the explosive.

Rate of detonation = rate of detonation, in meters per second, of the explosive compressed at the density "d."

\* Indicates decomposition in accordance with the equation:  $C_3N_{12} = 3C + 6N_2$ .

† Is based on products actually observed:  $C_4N_{12} = 1.27C_2N_2 + 4.79N_2 + 0.46C$  (amorphous).



quantity of substance used, Wöhler and Martin obtained a series of curves asymptotic to the weight axis, from which the true explosion temperature of the compound was determined. Their results are given in table 8 and compared with standard values for the corresponding fulminates. There appears to be no regularity in the temperature of explosion. The azides of silver, cadmium, and lead detonate at a high temperature, while the azides of cobalt and manganese detonate at a much lower temperature. It should be pointed out that protracted heating at temperatures even

TABLE 8  
*Detonation temperatures of azides and fulminates*

AZIDES AND FULMINATES	DETONATION TEMPERATURES OF	
	Azides	Fulminates
	<i>degrees C.</i>	<i>degrees C.</i>
Silver.....	297	170
Lead.....	327	—
Mercurous.....	281	215
Cadmium.....	291	—
Zinc.....	289	205
Cuprous.....	174	—
Nickel*.....	200	—
Cobalt†.....	148	—
Manganese.....	203	—
Barium.....	152	—
Strontium.....	169	—
Calcium.....	158	—
Lithium.....	245	—
Thallium.....	320	—
Mercuric.....	—	215

\* Hydrated salt.

† Contained 13 per cent water.

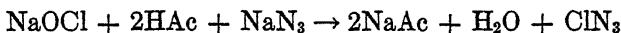
considerably lower than those noted in the table results in decomposition of these compounds (see page 205).

Wöhler and Martin also give data for the sensitivity of these azides to impact. Their results show that no apparent regularity exists between weights of substance employed and their sensitivity to shock. Some azides, such as mercurous, silver, and lead azides, become more sensitive as the quantity of substance is increased. In the case of nickel and cobalt azides the sensitivity increases up to a certain weight and then falls off again as the quantity of material is increased. Zinc and manganous azides, on the other hand, become less sensitive as the amount of substance is increased.

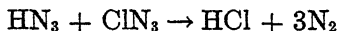
## VIII. NON-METALLIC AZIDES

The halogenoid (248) character of the azide radical is further substantiated by the fact that, like the halogens, it (1) forms compounds with other halogen or halogenoid groups and (2) replaces the halogens to form inorganic acyl azides. The few known members of both groups are described below.

*Chlorine azide*,  $\text{ClN}_3$  (184), is a colorless gas with a sweetish odor which is obtained by acidification, with acetic or boric acid, of a mixture of sodium hypochlorite and sodium azide in aqueous solution.



It dissolves in sodium hydroxide to give the hypochlorite and azide, and reacts with hydrazoic acid to give nitrogen and hydrochloric acid (183).



It is a tremendously explosive substance. It decomposes at  $400^\circ\text{C}$ . at 2 mm. pressure into the elements (110).

*Bromine azide*,  $\text{BrN}_3$  (209), may be prepared by passing bromine vapor over absolutely dry sodium azide. It also forms when silver and sodium azides are treated with bromine dissolved in ether, benzene, or ligroin. It is a mobile, very volatile, orange-red liquid which changes to a dark red solid at about  $-45^\circ\text{C}$ . It is instantly hydrolyzed, giving a mixture of hydrazoic and hypobromous acids which subsequently react to give nitrogen. It liberates iodine from a potassium iodide solution in quantities equivalent to the hypobromous acid formed.

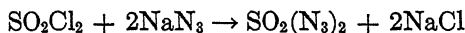
*Iodine azide*,  $\text{IN}_3$  (124), is obtained as a light yellow solid by interaction of an aqueous suspension of silver azide with iodine in ether or benzene. It is somewhat soluble in water and is decomposed largely even at  $0^\circ\text{C}$ . into the elements. It is hydrolyzed by alkaline solutions to give the azide, iodide, and iodate. With silver azide iodine azide gives molecular nitrogen, not the free dimeric azide radical as had been hoped. Metallic sodium yields sodium iodide and sodium azide; in insufficient quantity it yields sodium azide and free iodine, demonstrating that the azide radical is more electronegative than iodine. Iodine azide (as well as chlorine azide) is immediately decolorized by a sodium arsenite solution; sodium arsenate and sodium azide are formed (119, 120).

*Dicyandiazide*,  $(\text{CNN}_3)_2$ , was first isolated by Darzens (76), who believed the compound obtained by the action of cyanogen bromide upon sodium azide to be the monomer, cyanogen azide or carbon pernitride. Hart (127) showed that this could hardly be the case (although no molecular

weight determinations were made), since the compound reacted with ammonia in ethereal solution to give *dicyanamidazide*,  $\text{NCN}(\text{NH}_2)\text{N}_3$ , and ammonium azide. Dicyandiazide hydrolyzes in aqueous solution to yield carbon dioxide, hydrazoic acid, and cyanamide. It melts at  $40.3^\circ\text{C}$ ., begins to decompose at  $70^\circ\text{C}$ ., and explodes violently above  $170^\circ\text{C}$ . It is soluble in water, alcohol, ether, and chloroform, but insoluble in petroleum ether.

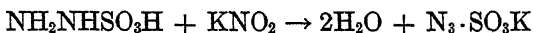
*Cyanuric triazide*,  $(\text{CN})_3(\text{N}_3)_3$ , may be obtained either from cyanuric chloride and sodium azide (175), or by diazotization of the corresponding hydrazide (96). The compound crystallizes in the hexagonal system (221), melts at  $94^\circ\text{C}$ ., and explodes when heated above  $170^\circ\text{C}$ . It has been prominently mentioned as a possible detonant (see p. 210).

*Sulfuryl azide*,  $\text{SO}_2\text{N}_6$ , was first prepared by Curtius and Schmidt (73) by the interaction of sulfuryl chloride and finely divided, slightly moist sodium azide.



After the reaction has gone to completion, the sodium chloride is dissolved out by the addition of ice water and the azide settles to the bottom as a water-white liquid. It is separated from the supernatant liquid and dried over anhydrous sodium sulfate. It explodes violently when heated, and often spontaneously at room temperatures. It has a suffocating odor and the pronounced physiological after-effects of hydrogen azide. It is hydrolyzed slowly in the cold by alcohol and water, but more quickly on warming. Alcoholic silver nitrate reacts with it instantaneously with the formation of silver azide. When heated in certain organic aromatic hydrocarbons, such as *p*-xylene, it undergoes decomposition with transient formation of the  $\left[ \text{>NSO}_2\text{N} < \right]$  residue which then breaks down further with liberation of sulfur dioxide.

Related to sulfuryl azide are the salts of *azidosulfonic acid*,  $\text{HSO}_3\text{N}_3$ , the analog of chlorosulfonic acid. Potassium azidosulfonate,  $\text{KSO}_3\text{N}_3$  (239), may be prepared by the treatment of a concentrated aqueous solution of potassium nitrite with finely powdered hydrazinesulfonic acid.



It crystallizes from solution in the form of large flat prisms which explode on heating. The addition of mineral acids probably causes the intermediate formation of azidosulfonic acid,  $\text{HSO}_3\text{N}_3$ , since hydrolysis of the product yields hydrazoic and sulfuric acids. The ammonium, sodium, and barium salts have been prepared, but their properties are not described.

*Carbonylazide*,  $\text{CO}(\text{N}_3)_2$ , is prepared by diazotization of carbohydrazide

dihydrochloride with sodium nitrite (69, 70). It is an extremely explosive, easily volatile, crystalline solid, very soluble in water, alcohol, and ether, but insoluble in petroleum ether. It undergoes hydrolysis to yield carbon dioxide and hydrazoic acid. Like sulfuryl azide, it decomposes in such solvents as benzene and xylene with probable transient formation of the  $\text{CO}(\text{N} \begin{smallmatrix} \diagup \\ \diagdown \end{smallmatrix})_2$  residue (67).

#### IX. ACKNOWLEDGMENTS

The author wishes to express his gratitude to Professor W. Conrad Fernelius of the Ohio State University and to Dr. John C. Bailar, Jr., of the University of Illinois for their kindly help and constructive criticisms in the preparation of the manuscript.

The author is also deeply indebted to Dr. L. V. Clark, formerly of the Bureau of Mines Experiment Station, Pittsburgh, now Industrial Fellow at the Mellon Institute for Industrial Research, for making available data from his personal files dealing with the use of azides as detonants and for his aid in reviewing that portion of this paper in which the explosive properties of the azides are discussed.

To Dr. W. E. Lawson, Director of the Eastern Laboratory of E. I. du Pont de Nemours and Company, the writer wishes to express his appreciation for use of an extensive patent bibliography dealing with the manufacture and application of azides as detonants.

In dedicating this article to Professor Arthur Wesley Browne of Cornell University the author has taken this opportunity to acknowledge not only his own indebtedness, but that of many other former students, to his and their teacher, to an eminent educator, and to an outstanding investigator in the field of the nitrogen compounds.

#### X. REFERENCES

- (1) ANGELI: *Atti accad. Lincei* [6] **5**, 732 (1927); *C. A.* **21**, 3603.
- (2) ANGELI: *Atti accad. Lincei* [5] **2**, I, 569; *Chem. Zentr.* **1893**, II, 559.
- (3) ANT-VUORINEN: *Suomen Kemistilehti* **4**, 89 (1931); *Chem. Zentr.* **1932**, II, 1148.
- (4) ARENS AND EGGERT: *Phot. Korr.* **67**, 17 (1931); *Chem. Zentr.* **1931**, II, 1537.
- (5) AUDRIETH: *J. Phys. Chem.* **34**, 538 (1930).
- (6) AUDRIETH: *J. Chem. Education* **7**, 2055 (1930).
- (7) AUDRIETH AND BROWNE: *J. Am. Chem. Soc.* **52**, 2799 (1930).
- (8) AUDRIETH, SMITH, G. B. L., AND BROWNE: *J. Am. Chem. Soc.* **49**, 2129 (1927).
- (9) BACH: *Z. physik. Chem.* **9**, 241 (1892).
- (10) BAMBERGER: *Ber.* **35**, 3893 (1902).
- (11) BECKMAN AND DICKINSON: *J. Am. Chem. Soc.* **50**, 1870 (1928).
- (12) BECKMAN AND DICKINSON: *J. Am. Chem. Soc.* **52**, 124 (1930).
- (13) BEKK: *J. Chem. Soc.* **108**, 200 (1915).
- (14) BEKK: *Z. wiss. Phot.* **14**, 105 (1914).

- (15) BENRATH AND RULAND: *Z. anorg. allgem. Chem.* **114**, 267 (1920).
- (16) BERGMAN AND SCHÜTZ: *Nature* **128**, 1077 (1931).
- (17) BERGMAN AND SCHÜTZ: *Z. physik. Chem.* **19B**, 389 (1932).
- (18) BERTHELOT AND MATIGNON: *Compt. rend.* **113**, 672 (1891).
- (19) BERTHELOT AND VIEILLE: *Bull. soc. chim.* [3] **11**, 744 (1894).
- (20) BIEHLER: *Knolls Mitt.* **1927**, 21; *C. A.* **24**, 659.
- (21) BIEHLER: *Arch. exptl. Path. Pharmacol.* **126**, 1 (1927); *C. A.* **22**, 823.
- (22) BIEHLER: *Ber. ges. Physiol. exptl. Pharmacol.* **49**, 829; *C. A.* **24**, 659.
- (23) BIRCKENBACH: *Z. anorg. allgem. Chem.* **214**, 94 (1933).
- (24) BIRCKENBACH AND KELLERMANN: *Ber.* **58**, 786 (1925).
- (25) BIRCKENBACH AND KELLERMANN: *Chem. Zentr.* **1926**, I, 1395.
- (26) BIRCKENBACH AND RÖRIG: *Chem. Zentr.* **1926**, I, 1492.
- (27) BLAU: *U. S. patent* 1,269,520.
- (28) BLAU AND WAMBACHER: *Monatsh.* **61**, 99 (1932).
- (29) BLECHTA: *Chimie & industrie*, Special No. 921 (June, 1933); *C. A.* **28**, 646.
- (30) BOOTH: *Science* **54**, 17 (1921).
- (31) BRINER AND WINKLER: *J. chim. phys.* **20**, 201 (1923).
- (32) BRINER AND WINKLER: *Helv. Chim. Acta* **6**, 429 (1923).
- (33) BRITTON AND ROBINSON: *Trans. Faraday Soc.* **28**, 531 (1932).
- (34) BROCKWAY AND PAULING: *Proc. Nat. Acad. Sci.* **19**, 860 (1933).
- (35) BROWNE: *J. Am. Chem. Soc.* **27**, 551 (1905).
- (36) BROWNE AND AUDRIETH: *J. Am. Chem. Soc.* **49**, 917 (1927).
- (37) BROWNE AND AUDRIETH: Unpublished investigations.
- (38) BROWNE AND VON HAZMBURG: *J. Am. Chem. Soc.* **48**, 2383 (1926).
- (39) BROWNE AND HOEL: *J. Am. Chem. Soc.* **44**, 2106 (1922).
- (40) BROWNE AND HOEL: *J. Am. Chem. Soc.* **44**, 2116 (1922).
- (41) BROWNE AND HOEL: *J. Am. Chem. Soc.* **44**, 2315 (1922).
- (42) BROWNE, HOEL, SMITH, G. B. L., AND SWEZEY: *J. Am. Chem. Soc.* **45**, 2541 (1923).
- (43) BROWNE AND HOLMES: *J. Am. Chem. Soc.* **35**, 672 (1913).
- (44) BROWNE, HOLMES, AND KING: *J. Am. Chem. Soc.* **41**, 1769 (1919).
- (45) BROWNE AND HOULEHAN: *J. Am. Chem. Soc.* **33**, 1742 (1911).
- (46) BROWNE AND LUNDELL: *J. Am. Chem. Soc.* **31**, 435 (1909).
- (47) BROWNE AND OVERMAN: *Z. anorg. allgem. Chem.* **94**, 217 (1916).
- (48) BROWNE AND SHETTERLEY: *J. Am. Chem. Soc.* **29**, 1305 (1907).
- (49) BROWNE AND SHETTERLEY: *J. Am. Chem. Soc.* **30**, 53 (1908).
- (50) BROWNE AND SHETTERLEY: *J. Am. Chem. Soc.* **31**, 221 (1909).
- (51) BROWNE AND SHETTERLEY: *J. Am. Chem. Soc.* **31**, 783 (1909).
- (52) BROWNE AND SMITH, G. B. L.: *J. Am. Chem. Soc.* **47**, 2698 (1925).
- (53) BROWNE AND WILCOXON: *J. Am. Chem. Soc.* **48**, 682 (1926).
- (54) BÜSSEM, ROSBAUD, AND GÜNTHER: *Z. physik. Chem.* **15B**, 58 (1931).
- (55) CLUSIUS: *Z. anorg. allgem. Chem.* **194**, 47 (1930).
- (56) COOKE: *Proc. Chem. Soc.* **19**, 213 (1904).
- (57) COPEMAN: *J. S. African Chem. Inst.* **10**, 18 (1927).
- (58) CRANSTON AND LIVINGSTONE: *J. Chem. Soc.* **1926**, 501.
- (59) CSERNECZKY: *Z. ges. Schiess-Sprengstoffw.* **24**, 169 (1928).
- (60) CURRIER AND BROWNE: *J. Am. Chem. Soc.* **44**, 2849 (1922).
- (61) CURTIUS: *Ber.* **23**, 3023 (1890).
- (62) CURTIUS: *Ber.* **24**, 3341 (1891).
- (63) CURTIUS: *Ber.* **26**, 1263 (1893).
- (64) CURTIUS: *Ber.* **26**, 1263 (1893).

- (65) CURTIUS: Ber. **29**, 759 (1896).
- (66) CURTIUS: Bull. soc. chim. **14**, 1004 (1895).
- (67) CURTIUS AND BERTHO: Ber. **59**, 565 (1926).
- (68) CURTIUS AND DARAPSKY: J. prakt. Chem. [2] **61**, 408 (1900).
- (69) CURTIUS AND HEIDENREICH: Ber. **27**, 2684 (1894).
- (70) CURTIUS AND HEIDENREICH: J. prakt. Chem. [2] **52**, 454 (1895).
- (71) CURTIUS AND RADENHAUSEN: J. prakt. Chem. [2] **43**, 207 (1891).
- (72) CURTIUS AND RISSOM: J. prakt. Chem. [2] **58**, 261 (1898).
- (73) CURTIUS AND SCHMIDT: Ber. **55**, 1571 (1922).
- (74) CUY, ROSENBERG, AND BRAY: J. Am. Chem. Soc. **46**, 1796 (1924).
- (75) DARAPSKY: Ber. **40**, 3033 (1907).
- (76) DARZENS: Compt. rend. **154**, 1232 (1912).
- (77) DEBOER, CLAUSING, AND ZECHER: Z. anorg. allgem. Chem. **160**, 129 (1927).
- (78) DENNIS: J. Am. Chem. Soc. **18**, 947 (1896).
- (79) DENNIS AND BENEDICT: J. Am. Chem. Soc. **20**, 225 (1898); Z. anorg. allgem. Chem. **17**, 25 (1898).
- (80) DENNIS AND BROWNE: J. Am. Chem. Soc. **26**, 577 (1904).
- (81) DENNIS AND BROWNE: Z. anorg. allgem. Chem. **40**, 68 (1904).
- (82) DENNIS AND DOAN: J. Am. Chem. Soc. **18**, 970 (1896).
- (83) DENNIS AND ISHAM: J. Am. Chem. Soc. **29**, 18 (1907).
- (84) DENNIS AND ISHAM: J. Am. Chem. Soc. **29**, 216 (1907).
- (85) DENNIS AND KORTRIGHT: Z. anorg. allgem. Chem. **6**, 35 (1894).
- (86) DENNSTEDT AND GÖHLICH: Chem.-Ztg. **21**, 876 (1897).
- (87) DIELS AND AUBART: Ann. **429**, 28 (1922).
- (88) DIMROTH: Ber. **40**, 2376 (1907).
- (89) DIMROTH AND PFISTER: Ber. **43**, 2757 (1910).
- (90) DRESSER AND BROWNE: J. Am. Chem. Soc. **53**, 4235 (1931).
- (91) DRESSER AND BROWNE: J. Am. Chem. Soc. **55**, 1963 (1933).
- (92) EBELER: Ber. **43**, 2613 (1910).
- (93) EINIG: German patent 435,654; Chem. Zentr. **1926**, II, 3072.
- (94) ERLÉNMEYER AND LEO: Helv. Chim. Acta **16**, 897 (1933).
- (95) FEIGL AND CHARGAFF: Z. anal. Chem. **74**, 376 (1928).
- (96) FINGER: J. prakt. Chem. [2] **75**, 103 (1907).
- (97) FISCHER: Ann. **190**, 67 (1878).
- (98) FRANK: German patent 529,601; Chem. Zentr. **1931**, II, 1728.
- (99) FRANKLIN: J. Am. Chem. Soc. **46**, 2137 (1924).
- (100) FRANKLIN: J. Am. Chem. Soc. **56**, 568 (1934).
- (101) FRANKLIN: J. Phys. Chem. **27**, 172 (1923).
- (102) FRANKLIN: Proc. 8th Int. Congr. Appl. Chem. [8] **6**, 119 (1912).
- (103) FRANKLIN: Science **56**, 28 (1922).
- (104) FRIEDLANDER: J. Am. Chem. Soc. **40**, 1945 (1918).
- (105) FRIEDRICH: Z. anorg. allgem. Chem. **84**, 389 (1914).
- (106) FROST, COTHMAN, AND BROWNE: J. Am. Chem. Soc. **55**, 3516 (1933).
- (107) GARDNER AND BROWNE: J. Am. Chem. Soc. **49**, 2759 (1927).
- (108) GARNER AND GOMM: J. Chem. Soc. **1931**, 2123.
- (109) GARNER AND MOON: Nature **131**, 513 (1933).
- (110) GLEU: Z. Physik **38**, 176 (1926).
- (111) GLEU: Ber. **61**, 702 (1928).
- (112) GLEU AND ROELL: Z. anorg. allgem. Chem. **179**, 233 (1929).
- (113) GRIESS: Ann. **137**, 39 (1866).
- (114) GROTTA: Ind. Eng. Chem. **17**, 134 (1925).

- (115) GÜNTHER, ANDREEW, AND RINGBOM: *Z. Elektrochem.* **36**, 211 (1930).
- (116) GÜNTHER, LEPIN, AND ANDREEW: *Z. Elektrochem.* **36**, 218 (1930).
- (117) GÜNTHER AND PERSCKE: *J. Chem. Soc.* **1930**, 100.
- (118) GÜNTHER, PORGER, AND ROSBAUD: *Z. physik. Chem.* **6B**, 459 (1929).
- (119) GUTMANN: *Ber.* **57**, 1956 (1924).
- (120) GUTMANN: *Z. anal. Chem.* **65**, 246 (1924).
- (121) HAHN: *Ber.* **65**, 64 (1932).
- (122) HAID AND GLOBIG: *Z. ges. Schiess-Sprengstoffw.* **21**, 162, 181 (1926).
- (123) HALE AND REDFIELD: *J. Am. Chem. Soc.* **33**, 1353 (1911).
- (124) HANTZSCH: *Ber.* **33**, 522 (1900).
- (125) HANTZSCH: *Ber.* **63**, 1782 (1930).
- (126) HANTZSCH: *Ber.* **66**, 1349 (1933).
- (127) HART: *J. Am. Chem. Soc.* **50**, 1922 (1928).
- (128) HARVEY: *Trans. Faraday Soc.* **29**, 653 (1933).
- (129) HEMPHILL: Thesis, Colgate University, 1925.
- (130) HENDRICKS AND PAULING: *J. Am. Chem. Soc.* **47**, 2904 (1925).
- (131) VON HERZ: *Z. ges. Schiess-Sprengstoffw.* **28**, 141 (1933).
- (132) HITCH: *J. Am. Chem. Soc.* **40**, 1195 (1918).
- (133) HITTORF: *Z. physik. Chem.* **10**, 593 (1892).
- (134) HODGKINSON: British patent 128,014.
- (135) HODGKINSON: British patent 129,152.
- (136) HOFMANN, K. A., AND HOFMANN, U.: *Ber.* **59**, 2574 (1926).
- (137) HOFMANN, K. A., HOCK, AND KIRMREUTHER: *Ann.* **380**, 131 (1911).
- (138) HOTH AND PYL: *Z. angew. Chem.* **42**, 888 (1929).
- (139) JOANNIS: *Compt. rend.* **118**, 713 (1894).
- (140) JUSTI: *Ann. Physik* [5] **10**, 983 (1931).
- (141) KAST AND HAID: *Z. angew. Chem.* **38**, 43 (1925).
- (142) KIRK AND BROWNE: *J. Am. Chem. Soc.* **50**, 337 (1928).
- (143) KÖCHER: *Klin. Wochschr.* **9**, 2160 (1930); *C. A.* **25**, 1907.
- (144) KOMPPA AND VUORINEN: *Suomen Kemistilehti* **2**, 60 (1929); *Chem. Zentr.* **1931**, II, 2135.
- (145) LABRUTO AND RANDISI: *Ann. chim. applicata* **22**, 319 (1932); *C. A.* **26**, 5372.
- (146) LANGMUIR: *J. Am. Chem. Soc.* **41**, 1543 (1919).
- (147) LINDEMANN: *Ber.* **63**, 1246 (1930).
- (148) LINDEMANN AND GROGER: *Ber.* **63**, 715 (1930).
- (149) LINDEMANN AND THIELE: *Ber.* **61**, 1529 (1928).
- (150) LOEW: *Ber.* **24**, 2947 (1891).
- (151) LOWNDES: *Trans. Faraday Soc.* **16**, 128 (1921).
- (152) MCKINNEY: Private communication.
- (153) MCKINNEY: Thesis, Cornell University, 1923.
- (154) MARSHALL: *Explosives*, Vols. I and II (1917); Vol. III (1932); (a) Vol. I, p. 509; (b) Vol. III, p. 159; (c) Vol. I, p. 507. J. and A. Churchill, London.
- (155) MARTIN: *J. Am. Chem. Soc.* **49**, 2133 (1927).
- (156) MATTER: German patent 302,561; *Chem. Zentr.* **1920**, II, 715.
- (157) MEISSNER: British patent 369,529; *Chem. Zentr.* **1932**, II, 267; see also German patent 594,363 and U. S. patent 1,923,933.
- (158) MENDELEJEFF: *Ber.* **23**, 3464 (1890).
- (159) MILES: *J. Chem. Soc.* **1931**, 2532.
- (160) MOLDENHAUER AND MÖRTTIG: *Ber.* **62**, 1954 (1929).
- (161) MOLES: *Anales soc. españ. fis. quim.* **26**, 133 (1928).
- (162) MOLES: *J. chim. phys.* **16**, 401 (1918).

- (163) MÜLLER AND BROUS: *J. Chem. Physics* **1**, 482 (1933).  
(164) MÜLLER AND BROUS: *J. Am. Chem. Soc.* **53**, 2428 (1931).  
(165) MURAOUR: *Bull. soc. chim.* [4] **51**, 1152 (1932).  
(166) NÄGELI AND VOGT-MARKUS: *Helv. Chim. Acta* **15**, 60 (1932).  
(167) OESTERLIN: *Z. angew. Chem.* **45**, 536 (1932).  
(168) OLIVERI-MANDALA: *Gazz. chim. ital.* **46**, II, 137 (1916).  
(169) OLIVERI-MANDALA: *Gazz. chim. ital.* **51**, I, 138 (1921).  
(170) OLIVERI-MANDALA: *Gazz. chim. ital.* **52**, II, 139 (1922).  
(171) OLIVERI-MANDALA: *Gazz. chim. ital.* **59**, 699 (1929).  
(172) OLIVERI-MANDALA: *Gazz. chim. ital.* **60**, 878 (1930).  
(173) OLIVERI-MANDALA AND COMELLA: *Gazz. chim. ital.* **52**, I, 112 (1922).  
(174) OLIVERI-MANDALA AND IRRERA: *Gazz. chim. ital.* **62**, 34 (1932).  
(175) OTT AND OHSE: *Ber.* **54**, 179 (1921).  
(176) PAULING: *Z. physik. Chem.* **8B**, 326 (1930).  
(177) PERATONER AND ODDO: *Gazz. chim. ital.* **25**, II, 13 (1895).  
(178) PERATONER AND ODDO: *Gazz. chim. ital.* **30**, II, 95 (1900).  
(179) PERSCHKE: *Ber.* **62**, 3054 (1929).  
(180) PICCARD AND THOMAS: *Helv. Chim. Acta* **6**, 1039 (1923).  
(181) RAMSPERGER: *J. Am. Chem. Soc.* **51**, 2134 (1929).  
(182) RASCHIG: *Schwefel und Stickstoffstudien*. Verlag Chemie, Leipzig (1924).  
    (a) p. 205; (b) p. 213.  
(183) RASCHIG: *Z. angew. Chem.* **23**, 972 (1910).  
(184) RASCHIG: *Ber.* **41**, 4194 (1908).  
(185) RASCHIG: *Ber.* **48**, 2088 (1915).  
(186) RATHSBURG: *British patent* 188,302; *Chem. Zentr.* **1923**, II, 940.  
(187) REITH AND BOUWMAN: *Pharm. Weekblad* **67**, 475 (1930); *C. A.* **24**, 3969.  
(188) REMY-GENNETÉ: *Ann. chim.* [10] **19**, 263 (1932).  
(189) *Reports of the Progress of Applied Chemistry* (Soc. Chem. Ind.) **1919**, p. 565.  
(190) *Reports of the Progress of Applied Chemistry* (Soc. Chem. Ind.) **1920**, p. 563.  
(191) RICCA AND PIRRONI: *Gazz. chim. ital.* **59**, 379 (1929).  
(192) RICCA AND PIRRONI: *Gazz. chim. ital.* **59**, 564 (1929).  
(193) RIEGGER: *Thesis*, Cornell University, 1911.  
(194) RIEGGER: *J. Am. Chem. Soc.* **33**, 1569 (1911).  
(195) ROTH AND MÜLLER: *Ber.* **62**, 1188 (1929).  
(196) SCHMIDT: *Ber.* **57**, 704 (1924).  
(197) SHATTENFROH: *Arch. Hyg.* **27**, 231.  
(198) SIDGWICK: *J. Chem. Soc.* **1929**, 1108.  
(199) SIDGWICK: *Structure of the Organic Azides*. *Trans. Faraday Soc.*, Preprint (1934).  
(200) SIDGWICK: *Some Physical Properties of the Covalent Link in Chemistry*, pp. 194-8. Cornell University Press, Ithaca, N. Y. (1933).  
(201) SIDGWICK, SUTTON, L. E., AND THOMAS: *J. Chem. Soc.* **1933**, 406.  
(202) SKRAUPY: *German patent* 414,517; *Chem. Zentr.* **1925**, II, 1623.  
(203) SMITH, G. B. L., GROSS, BRANDES, AND BROWNE: *J. Am. Chem. Soc.* **56**, 1116 (1934).  
(204) SMITH, G. B. L., WARTTMAN, AND BROWNE: *J. Am. Chem. Soc.* **52**, 2806 (1930).  
(205) SMITH, G. B. L., WILCOXON, AND BROWNE: *J. Am. Chem. Soc.* **45**, 2604 (1923).  
(206) SOMMER: *Ber.* **48**, 1833 (1915).  
(207) SOMMER AND PINCAS: *Ber.* **48**, 1963 (1915).  
(208) SOMMER AND WEISE: *Z. anorg. allgem. Chem.* **94**, 51 (1916).  
(209) SPENCER: *J. Chem. Soc.* **127**, 217 (1925).



- (210) STAUDINGER: J. Soc. Chem. Ind. **33**, 693 (1910); German patent 273,667.  
(211) STERN: Klin. Wochschr. **6**, 304 (1927); C. A. **21**, 1847.  
(212) STETTbacher: Z. ges. Schiess-Sprengstoffw. **9**, 381, 391 (1914).  
(213) STETTbacher: Z. ges. Schiess-Sprengstoffw. **21**, 53, 69 (1926).  
(214) STETTbacher: Z. ges. Schiess-Sprengstoffw. **28**, 77 (1933).  
(215) STOLLE: Ber. **41**, 2811 (1908).  
(216) STOLLE: German patent 205,683; Chem. Zentr. **1909**, I, 804.  
(217) STOLLE: German patent 415,855; Chem. Zentr. **1925**, II, 1625.  
(218) STRECKER AND OXENIUS: Z. anorg. allgem. Chem. **218**, 151 (1934).  
(219) SUHRMANN AND CLUSIUS: Z. anorg. allgem. Chem. **152**, 52 (1926).  
(220) SUTTON, L. E.: Nature **128**, 639 (1931).  
(221) SUTTON, T. C.: Phil. Mag. [7] **15**, 1001 (1933).  
(222) SUTTON, T. C.: Nature **128**, 872 (1931).  
(223) SZARVASY: J. Chem. Soc. **77**, 603 (1900).  
(224) TANATAR: Ber. **32**, 1399 (1899).  
(225) TAYLOR AND COPE: Bur. Mines Tech. Paper No. 162 (1917).  
(226) TAYLOR AND RINKENBACK: Z. ges. Schiess-Sprengstoffw. **20**, 114 (1925); Army Ordinance **5**, 824 (1925).  
(227) TAYLOR AND RINKENBACK: Bur. Mines Repts. Investigations No. 2513 (1923); J. Franklin Inst. **196**, 551 (1923).  
(228) TAYLOR AND RINKENBACK: J. Franklin Inst. **204**, 369 (1927).  
(229) THIELE: Ber. **41**, 2681 (1908).  
(230) THIELE: Ber. **41**, 2806 (1908).  
(231) THIELE: Ber. **44**, 2522 (1911).  
(232) THIELE: Ann. **271**, 127 (1892).  
(233) THIELE: German patent 66,806.  
(234) THIELE AND OSBORNE: Ann. **305**, 80 (1899).  
(235) TIEDE: Ber. **49**, 1742 (1916).  
(236) TIEDE: German patent 417,508; Chem. Zentr. **1925**, II, 1790.  
(237) TIEDE AND DOMCKE: Ber. **46**, 4100 (1913).  
(238) TILDEN AND MILLAR: J. Chem. Soc. **63**, 257 (1893).  
(239) TRAUBE AND VOCKERODT: Ber. **47**, 938 (1914).  
(240) TURRENTINE: J. Am. Chem. Soc. **33**, 803 (1911).  
(241) TURRENTINE: J. Am. Chem. Soc. **34**, 375 (1912).  
(242) TURRENTINE: J. Am. Chem. Soc. **34**, 385 (1912).  
(243) TURRENTINE: J. Am. Chem. Soc. **36**, 23 (1914).  
(244) TURRENTINE: J. Am. Chem. Soc. **34**, 382 (1912).  
(245) USHER AND VENKATESWARAN: J. Chem. Soc. **115**, 613 (1919).  
(246) VOURNASOS: Z. anorg. allgem. Chem. **164**, 263 (1927).  
(247) VOURNASOS: Praktika **7**, 227 (1932); Chem. Zentr. **1933**, II, 522.  
(248) WALDEN AND AUDRIETH: Chem. Rev. **5**, 339 (1928).  
(249) WATTENBERG: Ber. **63**, 1667 (1930).  
(250) WERNER: Proc. Chem. Soc. **28**, 257 (1911).  
(251) WESENBERG: German patent 506,771; Chem. Zentr. **1930**, II, 2945.  
(252) WEST: J. Chem. Soc. **77**, 705 (1900).  
(253) WEYR: Z. ges. Schiess-Sprengstoffw. **24**, 218 (1929).  
(254) WIENHAUS AND ZIEHL: Ber. **65**, 1461 (1932).  
(255) WILCOXON AND GROTTA: U. S. patent 1,628,380; C. A. **21**, 2137.  
(256) WILCOXON, MCKINNEY, AND BROWNE: J. Am. Chem. Soc. **47**, 1916 (1925).  
(257) WILICENUS: Ber. **25**, 2084 (1892); German patent 66,813.  
(258) WÖHLER: Z. angew. Chem. **24**, 2089 (1911).

- (259) WÖHLER: *Z. angew. Chem.* **35**, 545 (1922).
- (260) WÖHLER: *Chem.-Ztg.* **35**, 1096 (1911).
- (261) WÖHLER AND KRUPKO: *Ber.* **46**, 2045 (1913).
- (262) WÖHLER AND MARTIN: *Z. angew. Chem.* **27**, 335 (1914).
- (263) WÖHLER AND MARTIN: *Ber.* **50**, 586 (1917).
- (264) AUDRIETH: *Z. physik. Chem.* **165**, 323 (1933).
- (265) BURLOT: *Compt. rend.* **197**, 1223 (1933).
- (266) HAID, GOETZE, SELLE, KOENEN, SCHMIDT, AND BECKER: *Jahrb. chem.-tech. Reichsanstalt* **8**, 102 (1931); *C. A.* **26**, 3669.
- (267) MICHEL-LEVY AND MURAOUR: *Compt. rend.* **198**, 825 (1934).
- (268) MICHEL-LEVY AND MURAOUR: *Compt. rend.* **198**, 1499 (1934).
- (269) MURAOUR AND TRILLAT: *Chimie & industrie* **30**, 39 (1933).
- (270) ROLLET AND WOHLGEMUTH: *Compt. rend.* **198**, 1772 (1934).
- (271) SCHUMACHER: *Compt. rend.* **197**, 917 (1934).
- (272) SUTTON, T. C.: *Nature* **133**, 463 (1934).
- (273) WÖHLER AND MARTIN: *Z. angew. Chem.* **30**, 33 (1917).
- (273a) WENDT AND GRUBB: *Science* **52**, 159 (1920); *Chem. met. Eng.* **22**, 771 (1920).

#### *Patents*

- (274) Preparation of lead azide.—LIGNOSE, A. G.: German patent 440,568, *Chem. Zentr.* **1927**, I, 1723. DARIER AND GOUDET: U. S. patent 1,349,411. FRIEDERICH: German patent 365,773, *Chem. Zentr.* **1923**, II, 781; U. S. patent 1,478,429. LOWNDES: British patent 160,953, *Chem. Zentr.* **1921**, IV, 213. MATTER: British patent 300,401, *Chem. Zentr.* **1929**, I, 1253; German patent 310,090, *Chem. Zentr.* **1920**, IV, 207; German patent 510,420; *C. A.* **25**, 967; German patent 513,933, *Chem. Zentr.* **1931**, I, 1053; French patent 532,610, *Chem. Zentr.* **1922**, II, 908; French patent 663,841, *Chem. Zentr.* **1929**, II, 2625. MEISSNER: British patent 359,659; *C. A.* **26**, 4920; German patent 514,012; *C. A.* **25**, 1341; French patent 702,415; *C. A.* **25**, 4405; U. S. patent 1,959,731. RINKENBACK: U. S. patent 1,914,530. RINTOUL: British patent 142,898. Soc. d'Études Chim.: German patent 349,584; *Chem. Zentr.* **1922**, II, 1095. Sprengluft Ges.: German patent 396,438; *Chem. Zentr.* **1924**, II, 1304.
- (275) Use of lead azide as detonant (or component thereof).—BLAZZI: U. S. patent 1,950,019. BUELL: British patent 21,082 (1914); U. S. patent 1,174,669. BURKHARD: British patent 16,405 (1914). CLAESSEN: British patent 13,086 (1913); British patent 16,456 (1913); German patent 265,025, *Chem. Zentr.* **1913**, II, 1445; German patent 281,497; German patent 284,400; German patent 310,049, *Chem. Zentr.* **1921**, IV, 447; French patent 451,925; French patent 459,979. COOK AND GROTTA: U. S. patent 1,385,245; 1,406,977. DYNAMIT, A. G.: British patent 2,407 (1912). ESCHBACH: British patent 156,429, *Chem. Zentr.* **1921**, II, 765; British patent 172,914; British patent 304,144; German patent 305,100, *Chem. Zentr.* **1921**, II, 885; German patent 341,960, *Chem. Zentr.* **1922**, II, 52; German patent 363,671; U. S. patent 1,438,431. ESCHBACH AND WIPPENHOLM: French patent 710,999, *Chem. Zentr.* **1931**, II, 3705. Étab. Davey, Bickford, Smith et Cie: British patent 196,593, *Chem. Zentr.* **1923**, IV, 306; German patent 412,651, *Chem. Zentr.* **1925**, II, 703. FRIEDERICH: British patent 138,083; British patent 180,605, *Chem. Zentr.* **1922**, IV, 974; German patent 308,539, *Chem. Zentr.* **1921**, IV, 1161; German patent 339,202, *Chem. Zentr.*

- 1921, IV, 728; German patent 358,367, Chem. Zentr. 1922, IV, 1155; German patent 359,426, Chem. Zentr. 1923, II, 503; German patent 369,104, Chem. Zentr. 1923, IV, 39; German patent 373,633, Chem. Zentr. 1923, IV, 175; French patent 543,762; U. S. patent 1,424,462; U. S. patent 1,552,836. GELM: U. S. patent 1,406,844. GROTTA: U. S. patent 1,453,976; U. S. patent 1,481,361. HARLE: U. S. patent 1,488,787. HERZ: British patent 187,012, Chem. Zentr. 1923, II, 781; German patent 362,432, Chem. Zentr. 1923, II, 371; German patent 443,551, Chem. Zentr. 1927, II, 531; U. S. patent 1,498,001. HUDSON: U. S. patent 1,329,525. HYRONIMUS: British patent 1,819 (1908); German patent 224,669, Chem. Zentr. 1910, II, 771; French patent 384,792; U. S. patent 908,674. Imperial Chemical Industries: British patent 362,048, Chem. Zentr. 1933, I, 3663. KOWASTCH: Canadian patent 222,375; U. S. patent 1,424,487. LEWIS: Canadian patent 340,569; C. A. 28, 4235. MATTER: Canadian patent 176,610; British patent 280,249, Chem. Zentr. 1928, I, 2559; British patent 303,975, Chem. Zentr. 1931, I, 2712; German patent 570,044, Chem. Zentr. 1933, I, 2632; U. S. patent 1,239,613; U. S. patent 1,240,236; U. S. patent 1,254,147. RATHSBURG: British patent 177,744, Chem. Zentr. 1922, IV, 810; British patent 185,555, Chem. Zentr. 1923, II, 370; U. S. patent 1,470,104. RATHSBURG AND FRIEDERICH: British patent 195,344; French patent 554,390, Chem. Zentr. 1924, I, 1302. Rheinisch-Westfälische Sprengstoff A. G.: German patent 238,942. RUNGE: Canadian patent 181,129; U. S. patent 1,168,746; U. S. patent 1,185,830. SMITH: U. S. patent 1,502,754. SNELLING: U. S. patent 1,353,805; U. S. patent 1,462,074; U. S. patent 1,462,075. Soc. d'Études Chim.: British patent 130,166. Sprengluft Ges.: British patent 152,335; German patent 379,940, Chem. Zentr. 1924, I, 386. Sprengstoffwerke Nahnsen: German patent 380,010, Chem. Zentr. 1923, IV, 903. STINE: U. S. patent 1,309,552; U. S. patent 1,313,650. SYMMES: U. S. patent 1,480,795. Werkzeugmaschinen Fabrik Oerlikon: British patent 309,114, Chem. Zentr. 1931, I, 1053. WÖHLER: British patent 4,468 (1908); German patent 196,824, Chem. Zentr. 1908, I, 1439; French patent 387,640; U. S. patent 904,289; U. S. patent 1,128,394.
- (276) Use of aluminum, aluminum alloy, iron, or paper containers for lead azide detonants.—BROWNSDON: British patent 243,771. ESCHBACH: British patent 204,277; C. A. 18, 905; British patent 151,572; C. A. 15, 599. ESCHBACH AND FRIEDERICH: German patent 420,012, Chem. Zentr. 1926, I, 1493; German patent 439,582, Chem. Zentr. 1927, I, 1400; German patent 443,727, Chem. Zentr. 1927, II, 531; German patent 519,707, Chem. Zentr. 1931, I, 2834. Étab. Davey, Bickford, Smith et Cie: British patent 210,341. HARLE: Canadian patent 256,564, Chem. Zentr. 1926, II, 2260. Ver. Aluminumwerke A. G.: German patent 532,624, Chem. Zentr. 1931, II, 3705. WEIR: British patent 211,179, Chem. Zentr. 1924, II, 786. ZÜNDRÜTCHEN: German patent 401,345, Chem. Zentr. 1924, II, 2512.
- (277) Preparation and use of cyanuric triazide.—ORTT: German patent 345,794; German patent 350,564, Chem. Zentr. 1922, II, 1194; German patent 355,926; C. A. 17, 1242; German patent 550,479; U. S. patent 1,390,378. Soc. Anon. Pour l'Ind.: British patent 170,359, Chem. Zentr. 1922, II, 1118; French patent 531,088, Chem. Zentr. 1922, II, 1118.
- (278) Preparation of mercurous azide.—BLECHTA: French patent 704,994, Chem. Zentr. 1932, I, 1325. GROTTA: Canadian patent 246,338, Chem. Zentr. 1926, I, 553; U. S. patent 1,439,099; U. S. patent 1,533,798.
- (279) Use of sodium azide in priming compositions.—BUELL: U. S. patent 1,184,316.

# PRINCIPLES OF AN ELECTRONIC THEORY OF ORGANIC REACTIONS

CHRISTOPHER K. INGOLD

*Department of Chemistry, University College, University of London, England*

*Received July 18, 1934*

## CONTENTS

I. Introduction: scope of the paper.....	225
II. Polarization and polarizability.....	226
A. Polarization and polarizability of molecules.....	226
B. Polarization and polarizability of groups.....	226
C. Electrical picture of molecular activation.....	227
D. Electrical picture of molecular stability.....	228
III. Electron displacement.....	230
A. Electronic theory in relation to structure.....	230
B. Inductive and electromeric effects.....	231
C. Mesomeric and inductomeric effects.....	232
D. Electron displacement mechanisms in states and activations.....	232
E. The external molecular field.....	233
F. The combination of polar effects.....	234
G. Analogies with preëlectronic theories.....	235
IV. Polar classification of atomic groups and systems.....	236
A. Genesis of the inductive effect.....	236
B. Induction and evidence of dipole moment measurements.....	239
C. Genesis of the mesomeric effect.....	241
D. Source of the energy of mesomerism (valency degeneracy).....	249
E. Mesomerism and the evidence of dipole moment measurements.....	253
F. Mesomerism and the evidence of infra-red spectroscopy.....	257
G. Mesomerism and the evidence of thermochemistry.....	260
H. Inductomeric polarizability.....	263
I. Electromeric polarizability.....	264
V. Polar classification of reagents.....	265
A. Generalization of the conception of oxidation-reduction.....	265
B. Nucleophilic reagents.....	266
C. Electrophilic reagents.....	268
D. Classification of some particular reagents.....	270

## I. INTRODUCTION: SCOPE OF THE PAPER

The electronic theory of organic reactions has now developed to a stage at which it seems desirable to give a connected statement of its principles. Actually these principles have emerged piece-meal in connection with various separate applications of the theory; but the presentation of

the subject on a corresponding plan has the disadvantage that the common viewpoint tends to become obscured in what at the outset constitutes a mass of undesired detail. This paper, which reproduces with some additions the initial part of a course of lectures delivered in Stanford University in 1932, is an attempt to set forth the ideas of the theory, and to show that they are consistent with other theories and with physical evidence, quite apart from the *a posteriori* justification which results from specific applications to empirical data within that field of chemistry to which the theory especially applies.

## II. POLARIZATION AND POLARIZABILITY

### *A. Polarization and polarizability of molecules*

The simplest molecular model which can serve as a basis for an electrical conception of reactions is that which visualizes a space-distribution of atomic nuclei and electrons, these being regarded as point charges, maintained by elastic forces about fixed relative positions. This simple picture is of considerable utility, notwithstanding that in the development of the theory it has to be elaborated by the introduction of the idea of states involving a statistically continuous distribution of electron density; but whether we accept the simpler or the more complex conception, it remains true that the specification of a molecule in terms of the electrical model requires a knowledge of two kinds of electrical quantities concerned respectively with the positions and the mobilities of the charges,—in other words, with the state of polarization and with the polarizability of the system of charges. The terms “polarization” and “polarizability” have quantitative meaning in relation to a molecule, since each may be measured by certain accurately definable magnitudes. The principal magnitude, which describes the gross intrinsic polarization of a molecule, is the dipole moment  $\mu$ , and the most important of the quantities which measure the polarizability is the mean coefficient of polarizability  $\alpha$ . These two quantities differ in dimensions by an amount corresponding to the dimensions of an electric force, and the total moment developed by a molecule in the presence of a force  $F$  is therefore the sum of two terms  $\mu_x + \alpha_x F_x$  for a given direction  $x$ .

### *B. Polarization and polarizability of groups*

It is necessary to inquire to what extent the terms “polarization” and “polarizability” can be given precise meaning when applied, not only to complete molecules, but also to atoms and atomic groups in combination. The question may be answered by noting the extent to which such quantities as  $\mu$  and  $\alpha$ , which measure the polarization and polarizability of com-

plete molecules, are susceptible of analysis into characteristic contributions attributable to atoms and atomic groups. J. J. Thomson (43) originally suggested that dipole moments should be calculable as the vector sums of characteristic group contributions, provided that these have known relative directions; and the experimental researches of K. Höjendahl, C. P. Smyth, and J. W. Williams have shown that a rough relationship of this kind certainly exists. On the other hand, all attempts to establish for dipole moments an *accurately* additive relationship have failed; this is intelligible, because an accurately additive behavior would have implied the absence of intramolecular electrical interaction, and such interaction is not only a necessary property of the molecular model, but also an indispensable mechanism in the theory of organic reactions. The actual situation is that dipole moments may be represented as an additive property subject to disturbance through interaction, and the numerous published tables of atomic and group moments need not be misleading if it is agreed first to ascribe to groups characteristic moments, and then to treat intramolecular interaction as a superimposed effect. A closely similar situation prevails with reference to the analysis of coefficients of polarizability, which are directly obtainable from molecular refractions after certain small corrections have been applied. It is well known that molecular refractions exhibit an approximation to an additive principle; and that this principle is subject to disturbances which, in the classical development of spectrochemistry, are partly admitted as constitutive influences, and partly obscured in the averaging processes involved in the calculation of atomic and group constants. Thus groups as well as molecules may be considered to possess a duplex electrical specification—a state of polarization and of polarizability—provided it is recognized that, owing to intramolecular interaction, the appropriate measuring magnitudes cannot be regarded as strictly constant and characteristic.

### *C. Electrical picture of molecular activation*

It is a commonplace of electronic theories that chemical change is an electrical transaction, and that reagents act by virtue of a constitutional affinity, either for electrons or for atomic nuclei. When, for example, an electron-seeking reagent attacks some center in an organic molecule, reaction will supervene provided that center is able to supply electrons to the requisite extent; and the development of that critical electronic density at the site of reaction is an essential feature of the development of the energy of activation. The more readily the necessary electron density can be furnished, the more the reaction will be facilitated; thus the mechanism of electron supply becomes the main consideration relating to constitutional polar effects on reaction velocity. For instance, if it is known

that the replacement of one group by another at some point other than the focus of reaction leads to an increase in reactivity in the sense of a decrease in critical energy and an increase in speed, then the electropolar interpretation would be that the second substituent is more effective than the first in supplying electrons to the reaction zone. A similar series of statements would apply to nucleus-seeking reagents, excepting naturally that the primary necessity is now a low electron density at the site of reaction, and that facilitating groups are therefore those which withdraw electrons from the reaction zone.

The electropolar effect of a group is to be regarded as determined not merely by its field in the resting state, but by the electrical influence which it exerts under the stress of the electronic demands of the reaction itself. The permanent state of polarization of the group will certainly contribute, but the deformation, which arises in consequence of its polarizability and the electrical field induced by the attacking reagent, will contribute also. Thus the total electrical effect of the group will have the character of a duplex quantity, in which suitable measures of the polarization and of the polarizability of the group, and of the electrical demands of the reaction, play parts more or less analogous in function to those of  $\mu$ ,  $\alpha$ , and  $F$ , respectively, in the binomial  $\mu + \alpha F$ . Controversies have arisen because two reactions, each considered by its investigator to be a test of "polarity," do not place groups in an identical sequence with regard to their powers of facilitating the reactions; obviously, however, the order in which groups arrange themselves with respect to their polarization will bear no simple relation to the order of their polarizability, and the relative importance of the two contributory effects, dependent respectively on these two independent electrical characteristics, must necessarily vary with the nature of the reaction.

#### *D. Electrical picture of molecular stability*

The study of constitutional effects on reactivity must include the consideration not only of reaction velocities, but also of chemical equilibria. For the discussion of the latter it is necessary to envisage the internal energy of constitution of molecules in their ordinary states from an electropolar point of view.

One aspect of this conception may be illustrated by considering two molecules  $A-B$  and  $A-B'$ , such that the group  $B'$  tends to repel electrons more than does the group  $B$  towards the common molecular residue  $A$ . Let it be supposed that the residue  $A$  undergoes some constitutional change, and that the altered group  $A_1$  has a slightly greater affinity for electrons than had the original group. Evidently the structure  $A-B'$  has more to gain (or less to lose) as regards stability on conversion into  $A_1-B'$

than had the structure  $A-B$  on transformation into  $A_1-B$ . If  $A-B$  and  $A-B'$  are related as factor and product respectively of a reversible reaction, the equilibrium between them will depend on the relationship between their free energies, and a similar statement will apply to the equilibrium between the factor  $A_1-B$  and its product  $A_1-B'$ . Clearly the electropolar effect of the constitutional change represented by the alteration of  $A$  to  $A_1$  will be to transfer thermodynamic stability from factor to product, and therefore to displace equilibrium in the direction favoring the product. An identical principle will apply if, for example,  $B$  or  $B'$  is electron-attracting, and the slightly different descriptions relating to these and other possible cases need not be detailed.

It should be noted, however, that the relationship indicated is expected to hold generally and in principle only if the polar effect of the constitutional change represented by the conversion of  $A$  into  $A_1$  is *small*. The reason for this is that, owing to intramolecular electrical interaction, group dipoles are not characteristic constants, and that if two groups have different polarizabilities, their states of polarization may be affected to different extents by the electrical character of the residues to which they are attached. Thus, in the above example,  $B'$  is assumed to repel electrons more than  $B$ ; but if  $B$  happened to be the more polarizable group, and if the modification of the common molecular residue  $A$  conferred upon it a sufficiently strong additional attraction for electrons, then the original difference between the states of polarization of  $B$  and  $B'$  might become inverted as a result of the conversion of  $A$  into  $A_1$ ; the effect on equilibrium could then be the opposite of that stated. This is another consequence of the duplex character of "polarity," which has frequently been overlooked in discussions to which it is relevant.

The element of molecular stability considered above is concerned with the mutual electrostatic accommodation of the linked groups, and it therefore depends essentially on the mean positions of the electrons. We may expect to observe a second element of stability which is connected with the facility of electronic interplay between the groups, and is therefore dependent on the effective mobilities of the electrons.

In illustration let us consider a molecule  $A-B$  having something of the character of an electrical insulator, and suppose that the exchange of  $B$  for  $B'$  alters the properties of the links in such a way that the molecule  $A-B'$  has more of the nature of an electrical conductor. Evidently this circumstance may permit of a levelling out of local potentials present in  $A$ , with a consequential gain in stability. Now let it be supposed that the group  $A$  is replaced by a group  $A_1$ , having more, or more pronounced, local potentials capable of undergoing this levelling. The gain in stability accompanying the replacement of  $B$  by  $B'$  will clearly be greater when the attached



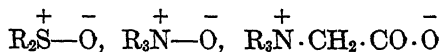
group is  $A_1$  than when it is A. Therefore, if  $A-B$  and  $A-B'$  are related as factor and product respectively of a reversible reaction, and if  $A_1-B$  and  $A_1-B'$  are similarly connected, the effect of the exchange of A for  $A_1$  will be to shift thermodynamic stability from the factor to the product, and therefore to displace the equilibrium in a corresponding direction. We cannot pursue this conception much further on the basis of the simple electronic model assumed at the outset, and its development must await the introduction of the quantum mechanical ideas considered later.

### III. ELECTRON DISPLACEMENT

#### A. *Electronic theory in relation to structure*

The electronic theory of valency (24) is now so celebrated that no explanation of its principles is necessary. Its adoption as a fundamental viewpoint in organic chemistry may be taken as axiomatic, and the interpretation of structure in terms of the theory of the duplet and the octet will be an accepted preliminary to the consideration of reactions. It may, however, be appropriate to refer to the conventions by means of which organic chemical structures are conveniently represented on the basis of the electronic theory.

First, the ordinary literal symbols for the elements are employed to denote atomic kernels, that is, atoms without the electrons of their valency shells. Secondly, the bond sign, either expressed, as in full graphic formulas, or implied, as in the customary abbreviated forms, is used to represent the shared electron-duplet or covalency. Thirdly, unshared electrons are not explicitly symbolized, since, if the valency shell is complete, their number is apparent from the number of covalencies. Fourthly, ionic centers are labelled  $+$  and  $-$  and free radical centers  $\cdot$ , since these labels not only emphasize important molecular conditions, but also remove any ambiguities which might otherwise arise in the presence of incomplete valency shells. Thus in the examples,



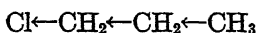
all octets are complete; but the signs serve to call attention to the circumstance that the corresponding atoms exhibit the covalencies which characterize them in unipolar sulfonium, ammonium, or oxide (e.g., hydroxide) ions (21), and that accordingly the compounds are to be regarded as dipolar ions.<sup>1</sup> On the other hand, the formulas  $R_3\overset{-}{C}$ ,  $R_3\overset{\cdot}{C}$ ,  $R_3\overset{+}{C}$  would be ambiguous but for the signs, which show that the first has two unshared

<sup>1</sup> With the notation suggested, the use of a special symbol (e.g.,  $\rightarrow$ ) when the ionic centers happen to be adjacent becomes unnecessary.

electrons, the second, one, and the third, none, and that only the first has a completed octet.<sup>2</sup> It may be added that the electrovalency, having no significance as a factor in molecular or ionic structure, requires no special symbol in organic chemistry, except in so far as the brackets [ ] of Werner's notation may be of convenience for distinguishing the individuals of a group of ions.

### *B. Inductive and electromeric effects*

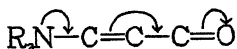
The significance of the electronic theory of valency as a basis for the development of an electrical conception of reactions is that it limits, by the principle requiring the preservation of stable electron groups, the forms of electron displacement which it is permissible to assume in the interpretation of intramolecular electrical interaction. There are two ways in which electron duplets may undergo displacement with preservation of atomic octets and without alteration of the original arrangement of atomic nuclei. The first method is characterized by the circumstance that the displaced electrons remain bound in their original atomic octets; displacements of this type were postulated (1923) by G. N. Lewis (25), who showed how the electrical dissymmetry arising from the unequal electron attractions of two linked atoms could be propagated along a molecule by a mechanism analogous to electrostatic induction. This mechanism was considered to illustrate a permanent molecular condition, that is, a state of polarization; it has been designated the *inductive* effect (11), and is represented by attaching to the bond sign an arrowhead indicating the direction towards which the electrons are concentrated. Thus the symbol



indicates that the electrical dissymmetry of the chlorine-carbon linking has caused the electrons of the carbon-carbon bonds to concentrate in the direction of the chlorine atom. The second method of electron displacement is characterized by the substitution of one duplet for another in the same atomic octet; displacements of this form were first assumed (1923-25) by T. M. Lowry (27), who showed how the entrance into an octet of an unshared duplet possessed by a neighboring atom could cause the ejection of another duplet, which would then either become unshared or initiate a similar exchange further along the molecule. This process was regarded as exemplifying, not a permanent molecular condition, but

<sup>2</sup> The labelling of atoms + and - with another significance than that of marking ionic centers tends to cause confusion. The notation  $\delta+$  and  $\delta-$  may be used without ambiguity to represent the acquisition of polarity through electron displacement, and N. V. Sidgwick's sign ( $\rightarrow$ ) for a molecular dipole (the crossed end being the positive end) is also useful.

an activation mechanism or polarizability effect; it has been termed the *electromeric* effect (16) and is represented by a curved arrow pointing from the duplet to the situation towards which displacement is assumed to occur. Thus the arrows contained in the symbol



connote duplet displacement in directions which could lead to the formation of the dipole,  $R_2\overset{+}{\text{N}}=\text{C}-\text{C}=\overset{-}{\text{C}}-\ddot{\text{O}}$ .

### C. Mesomeric and inductomeric effects

The idea of a permanent state of polarization associated with the same mode of electron displacement as that which characterizes the electromeric effect was introduced (1926) by C. K. Ingold and E. H. Ingold (16). According to this view the stable state of a system such as the one formulated just above is really intermediate between the formally non-polar and the dipolar forms: the displaced duplets which appear unshared in one formula or the other are considered to be partly shared in the stable state, whilst the displaced duplets which are represented as being shared in both forms are regarded as being permanently under the control, not of two atomic nuclei, but of three. This mode of polarization, originally designated "electronic strain," has been renamed the *mesomeric* effect, on account of considerations which will be indicated later. The complementary conception of a polarizability effect occurring in activation by a mechanism analogous to that of the inductive effect (15) is of later introduction into the theory (1933); it is termed the *inductomeric* effect.

### D. Electron displacement mechanisms in states and activations

The four effects enumerated in the two preceding paragraphs are distinguished partly by their electrical classification as polarizations or polarizabilities, and partly by their electronic mechanisms. It is convenient to be able to refer to these mechanisms as such, independently of whether they are being applied to a permanent state of the molecule or to an activation phenomenon. The process represented by the arrowheaded bond sign ( $\rightarrow$ ) is termed the *general inductive* mechanism (symbol *I*) and that denoted by the curved arrow ( $\curvearrowright$ ) is called the *tautomeric* mechanism (symbol *T*). The relationship between the four polar effects, their electrical classification, and their electronic mechanism, may be represented as in the following scheme:

ELECTRONIC MECHANISM	ELECTRICAL CLASSIFICATION	
	Polarization	Polarizability
General inductive ( $\rightarrow$ ) ( <i>I</i> ) Tautomeric ( $\curvearrowright$ ) ( <i>T</i> )	Inductive Mesomeric ( <i>M</i> )	Inductomeric Electromeric ( <i>E</i> )

It should be stated that there are certain special branches of the theory of organic chemical reactions which cannot adequately be dealt with solely on the basis of the above scheme. The two most important fields to which this statement applies are intramolecular rearrangement and free radicals. It will be evident that when allowance has to be made for possible redistributions of atomic nuclei amongst themselves, the types of electron displacement permitted by the octet theory will not be limited to those already described; thus the discussion of intramolecular rearrangements requires the introduction of additional special principles. Again, since the mechanisms of electron displacement already considered imply that all the electrons are paired, these mechanisms cannot be applied without extension to molecules which contain an unpaired electron; and thus some new principle is required for the adequate discussion of free radical chemistry. However, the necessity for certain additional theoretical principles in a few special cases does not seriously affect the generality of the picture already presented, and it still remains true that the mechanisms of electron displacement described above represent the central concept of the electronic theory of organic reactions.

#### *E. The external molecular field*

When we consider more closely the propagation of the general inductive effect, we find that the intensity of the effect reaching any point at some distance from the source of the disturbance must be partly dependent on the existence of an electric field in the space immediately outside the molecule. Thus in a chain of atoms  $X-C_\alpha-C_\beta-C_\gamma-C_\delta$  the  $X-C_\alpha$  dipole must affect the  $C_\beta-C_\gamma$  and  $C_\gamma-C_\delta$  linkings by direct action through space, as well as by relay through the intermediate links; and a similar statement applies to the effect of the  $C_\alpha-C_\beta$  dipole on the  $C_\gamma-C_\delta$  link. This kind of short-circuiting of the induction process was originally suggested by G. N. Lewis (25). The reason for assuming it arises from the circumstance, which is considered more fully and illustrated later, that the link dipoles become successively weaker with increasing distance along the chain, so that the contiguous ends of two successive link dipoles never neutralize each other. Each atom of the chain therefore acquires a residual charge, and consequently must of necessity exert an electrical field in the surrounding space.

The most important consequence of this partly external method of propagation of the general inductive effect is that the intensity with which it reaches any point in a molecule must depend not merely on the number of links which the effect has traversed, but also on various "air-line" distances (20), and on the spatial orientations of the links of the molecule relatively to the corresponding air-line distances (3). In certain types of molecules the external propagation of the general inductive effect becomes important, and the significance of air-line distances, and of spatial directions relatively to these distances, can be perceived in the reactivities exhibited. For this reason the externally propagated part of the general inductive effect is often referred to as the "direct effect," but the title is only one of convenience, and does not connote anything that can clearly be separated from the general inductive effect itself. Evidently the short-circuiting of induction depends essentially on the existence of the inductive mechanism and on its characteristic diminution with distance, and this short-circuiting must occur in molecules of all types and in all stages of the propagation of the effect. The "direct effect" is therefore best regarded as a special aspect or characteristic of the general inductive mechanism of electron displacement, and as applicable both to polarizations (inductive effect) and to polarizabilities (inductomeric effect) dependent on this mechanism.

#### *F. The combination of polar effects*

An important phase in the historical development of the subject was that concerned with the synthesis of the various electropolar effects into a coherent theoretical viewpoint. The first step in this direction was taken in 1924-25 by H. J. Lucas and his collaborators (28), who brought together the inductive and electromeric effects and showed how the former might be supposed to assist and give direction to the latter, as, for example, in the activation of olefins in their addition reactions:



The incorporation into the theory of the mesomeric and the inductomeric effects followed in the course of its development, as has already been indicated. Closely linked to the progressive association of these various ideas was the growth in the applicability of the theory to the problems of organic chemistry. Thus the very simple character of the mechanistic combination assumed by Lucas is to be associated with the definitely limited applicability of his theory. More elaborate combinations of polar effects were assumed in 1926 by R. Robinson and his collaborators (36), and by C. K. Ingold and E. H. Ingold (16). These authors achieved a

generalization sufficiently wide to permit an interpretation of the complex phenomena of aromatic substitution. Still, however, the theory was too incomplete to admit of satisfactory application to the problems of velocity and equilibrium in organic reactions as a whole; the necessary generalization has been one of the principal tasks before the writer and his colleagues since 1926.

### *G. Analogies with preëlectronic theories*

Although there is no close analogue in any preëlectronic theory of organic chemistry to the electronic theory of reactions considered as a whole, correspondencies may nevertheless be perceived between certain preëlectronic theories and some of the individual ideas which the electronic theory incorporates. Thus the early conception of A. Michael (30) and others, according to which groups possess an electropolar quality capable of "infecting" the adjacent parts of the molecule, has in common with the theory of general inductive electron displacement the idea of polar quality propagated by contact, independently of the precise constitutional details of the system. Similarly, J. Thiele's hypothesis of conjugation (42) shares with the theory of tautomeric electron displacement the conception of interaction between the centers of unsaturation separated by a single linking. A. Lapworth's theory (22) of molecular activations arising by a mechanism of valency redistribution of type analogous to that assumed by Thiele may be compared with the conception of the electromeric effect; the analogy is strengthened by the circumstance that Lapworth endowed his valencies with electropolar qualities, but this led him to the principle of alternate polarities, which is now known to be incorrect; it is also clear that a proper distinction between the functions of saturated and unsaturated centers was not maintained. B. Flürscheim's theory (8) of permanently redistributed valencies may be compared with the conception of the mesomeric effect. It was an important merit of Flürscheim's theory that it enabled its author to discuss not only problems of reactivity, but also those of chemical equilibrium, which are outside the scope of an activation theory; on the other hand, an inadequate appreciation of the distinctive functions of saturated and unsaturated atoms led to many valency distributions which would not now be approved, and the manner in which it was attempted to associate polar quality with the redistributed valencies is radically different from that adopted in the electronic theory of reactions. These preëlectronic theories do not form part of the foundations of the electronic theory of reactions in the sense that the former theories contain any acceptable ideas which are not expressed more definitely and with less associated error in the postulates already stated; on the other hand, it remains true that a number of the conceptions which the electronic theory

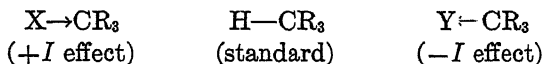
embodies have previously existed in organic chemistry, and that several partial collations of experimental data have been effected with their aid. It was of great assistance in the development of the electronic theory of reactions that these collations were available, and (sometimes after necessary corrections) could be incorporated in the larger generalization.

#### IV. POLAR CLASSIFICATION OF ATOMIC GROUPS AND SYSTEMS

##### *A. Genesis of the inductive effect*

The problem of classifying atomic groups and systems according to their polar effects is necessarily complex, owing to the dual character of polarity, and to the two methods of electron displacement permitted in the octet theory. We have, indeed, to consider separately the four polar effects distinguished in the preceding section. Of these four, only the two which represent polarization (the inductive and mesomeric effects) can be discussed solely on the basis of the constitution of the organic molecule and without reference to the character of the attacking reagent. These effects will be considered in the immediate sequel, and the inductive effect, which is the simpler and more general, will be taken first (11, 14, 25, 28).

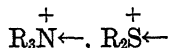
In the electronic theory of organic reactions inductive electron repulsions and attractions are always regarded as relative phenomena, and, by convention, the standard of reference is hydrogen. A group is said to repel electrons if it does so more than would hydrogen in the same molecular situation; thus a group X would correctly be described as electron-repelling in the compound  $X-CR_3$  if the electron density in the residue  $-CR_3$  were greater in this compound than in the compound  $H-CR_3$ . Similarly, the group Y is classified as electron-attracting in  $Y-CR_3$ , if the electron density in  $-CR_3$  is less in this compound than in the compound  $H-CR_3$ . Electron repulsions are conventionally distinguished by a positive sign, and electron attractions by a negative sign, so that the inductive effects illustrated may briefly be represented by the symbols  $+I$  and  $-I$ :



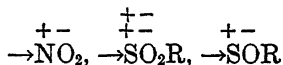
It has already been explained that the inductive effect exerted by groups must be influenced to some extent by molecular environment, and although this factor is usually of relatively small importance, it is convenient at the outset to reduce the problem of classification to a form in which the disturbance mentioned is minimal. This will be true if the groups considered are imagined to be present singly as substituents in derivatives of paraffins.

A major distinction must be drawn between charged and neutral groups. Anionic groups (e.g.,  $\bar{O}\rightarrow$ ,  $\bar{S}\rightarrow$ ) as a whole are expected to repel electrons

in comparison with neutral groups considered as a whole; similarly, cationic groups, e.g.,



as a class should attract electrons relatively to neutral groups as a class. The reason is that in anionic centers the atomic nuclei are overcompensated, and in cationic centers they are undercompensated, by the electrons of their own octets. Groups consisting of, or containing, dipolar ions, e.g.,



should attract electrons relatively to neutral groups, if the former are joined, as they usually are, through their cationic centers to the remainder of the molecule. It should be emphasized that these are general relationships, and that they take no account of the individual variations which may lead to occasional overlapping between the various series.

Individual distinctions depending on chemical types become apparent on considering a series of different groups in the same state of electrification. The matter may be illustrated by reference to neutral groups, although the arguments to be developed are, in principle, equally applicable to the grading of anionic, cationic, or dipolar groups amongst themselves. In the series  $\cdot CH_3$ ,  $\cdot NH_2$ ,  $\cdot OH$ ,  $\cdot F$ , the total nuclear and electronic charges are the same, but the distributed (protonic) portion of the nuclear charge in  $\cdot CH_3$  becomes progressively centralized in  $\cdot NH_2$ ,  $\cdot OH$ ,  $\cdot F$ , which therefore attract electrons successively more strongly in comparison with the methyl group (the polarity of alkyl groups is discussed below). Again, an atomic field (of either sign) is expected to suffer more loss through internal electronic deformation in larger than in smaller atoms, and thus electron attractions should diminish in the series  $\cdot F$ ,  $\cdot Cl$ ,  $\cdot Br$ ,  $\cdot I$ . A regular connection between the inductive effect of an atom and its position in the periodic table of the elements is thus adumbrated, and if R is a non-polar or feebly polar group (H, Alkyl) then all groups such as  $\cdot NR_2$ ,  $\cdot OR$ ,  $\cdot SR$ ,  $\cdot Hal$ ,  $\cdot NR$ ,  $\cdot O$ ,  $\cdot N$ , should attract electrons relatively to the methyl group. The extent of the attraction should increase with the number of the periodic group, and decrease with increasing number of the periodic series; it should also be greater in multiply-linked atoms than in corresponding singly-linked atoms.

It can be deduced from the premises stated above that the inductive effect of all alkyl groups is zero. This conclusion, which is consistent with the relationship already indicated between the inductive effect and the periodic classification of the elements, follows formally for the methyl group from the symmetry and non-polar character of methane and ethane,



and, for other alkyl groups, from analogous comparisons. It is, however, to be correlated with the circumstance that the paraffin molecular framework has been selected in order to standardize the effects of molecular environment. The real conclusion concerning alkyl groups is therefore that, unlike the groups already considered, the intrinsic polarity of which renders their classification at least qualitatively insensitive to the disturbance of molecular environment, alkyl groups will merely exert those polar effects which are impressed on them by the other groups present in the molecule. The important property of alkyl groups is that they are more

TABLE 1  
*Inductive effect*

---

Electron repulsion (+I)	
$\cdot\bar{\text{N}}\text{R} > \cdot\bar{\text{O}};$	
$\cdot\bar{\text{O}} > \cdot\bar{\text{S}} > \text{Se};$	
$(\text{CH}_3)_3\text{C}\cdot > (\text{CH}_3)_2\text{CH}\cdot > \text{CH}_3\cdot\text{CH}_2\cdot > \text{CH}_3\cdot$	
Electron attraction (−I)	
$^+\cdot\text{OR}_2 > ^+\cdot\text{NR}_3;$	
$^+\cdot\text{NR}_3 > ^+\cdot\text{PR}_3 > ^+\cdot\text{AsR}_3 > ^+\cdot\text{SbR}_3;$	
$^+\cdot\text{OR}_2 > ^+\cdot\text{SR}_2 > ^+\cdot\text{SeR}_2 > ^+\cdot\text{TeR}_2;$	
$^+\cdot\text{NR}_3 > ^{+-}\cdot\text{NO}_2; ^{+-}\cdot\text{SO}_2\text{R} > ^{+-}\cdot\text{SOR}; ^{+-}\cdot\text{SO}_2\text{R} > ^{+-}\cdot\text{SO}_3;$	
$^+\cdot\text{NR}_3 > ^+\cdot\text{NR}_2; ^+\cdot\text{OR}_2 > ^+\cdot\text{OR}; ^+\cdot\text{SR}_2 > ^+\cdot\text{SR};$	
$\cdot\text{F} > \cdot\text{OR} > \cdot\text{NR}_2 (> \cdot\text{CR}_3);$	
$\cdot\text{F} > \cdot\text{Cl} > \cdot\text{Br} > \cdot\text{I};$	
$\cdot\text{O} > \cdot\text{NR} > \cdot\text{CR}_2; \cdot\text{N} > \cdot\text{CR};$	
$\cdot\text{O} > \cdot\text{OR}; \cdot\text{N} > \cdot\text{NR} > \cdot\text{NR}_2;$	
$\cdot\text{C}:\text{CR} > \cdot\text{CR}:\text{CR}_2 (> \cdot\text{CR}_2\cdot\text{CR}_3)$	

---

*polarizable* than hydrogen, and thus  $\cdot\text{CH}_3$ , although non-polar when  $\text{CH}_3\text{—CH}_3$  is compared with  $\text{H—CH}_3$ , becomes an electron-repelling group when the comparison is between  $\text{CH}_3\text{→CO}_2\text{H}$  and  $\text{H—CO}_2\text{H}$ , or between  $\text{CH}_3\text{→CH}_2\text{Cl}$  and  $\text{H—CH}_2\text{Cl}$ . Since the majority of substituents commonly encountered are attractors of electrons, alkyl groups generally function as feebly electron-repelling groups; the possibility of the other type of behavior can, however, be foreseen, although unambiguous illustrations are difficult to find.

A case not covered by the preceding survey is that of unsaturated hydrocarbon radicals, such as ethenyl and phenyl. It has been empirically

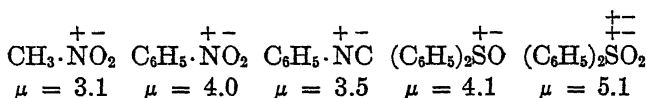
observed that alkyl groups behave as if they repel electrons when attached to such unsaturated systems, and it is probable that this relationship between saturated and unsaturated hydrocarbon radicals in combination is quite general (e.g.,  $\text{CH}_3 \rightarrow \text{C}_6\text{H}_5$ ). An equivalent statement, more consistent with the viewpoint set forth above, is that radicals such as ethenyl and phenyl are to be regarded as intrinsic attractors of electrons. The acetylenyl radical attracts electrons more strongly still. The theory of these effects is at present obscure.

The preceding conclusions are summarized for convenience in table 1. It will be appreciated that the whole argument is based on the hypothesis that the source of the energy of the inductive effect originating in any atom is the stability of its octet, including the effect thereon of nuclear charge and of electron screening.

### *B. Induction and the evidence of dipole moment measurements*

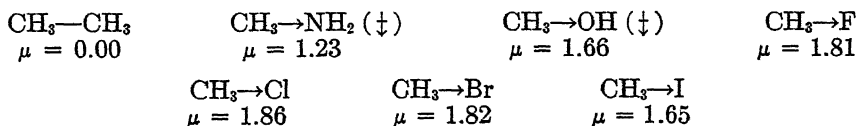
The significance of molecular dipole moments in relation to the theory of organic reactions was pointed out by C. K. Ingold in 1926 when few measurements of dipole moments were available (12). Since that time a large number of determinations have been made, and a certain amount of the argument of the preceding section can now be illustrated by their aid. The figures quoted are in Debye units ( $10^{-18}$  electrostatic unit).

One of the most obvious generalizations which the data reveal is that groups containing dipolar ions invariably confer on the molecule a large dipole moment (3 to 4 units). Not many illustrations amongst aliphatic substitution products are available, but the value for nitromethane has been determined, and there are a number of examples in the aromatic series:



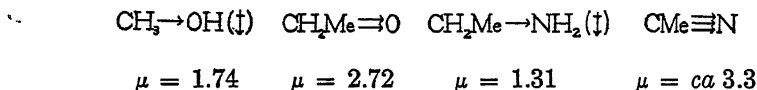
It is true that the aromatic illustrations involve an environmental (mesomeric) effect, but the magnitude of this (up to about 1 unit) is not such as to disturb the general conclusion. (Proof of the dipolar ionic structure of the isonitriles,  $\overset{-}{\text{C}} \cdot \overset{+}{\text{NR}}$ , has been given by N. V. Sidgwick and his collaborators (38).)

The effect on molecular dipole moments of the introduction of non-ionic substituents into saturated carbon systems may be exemplified by the following values:

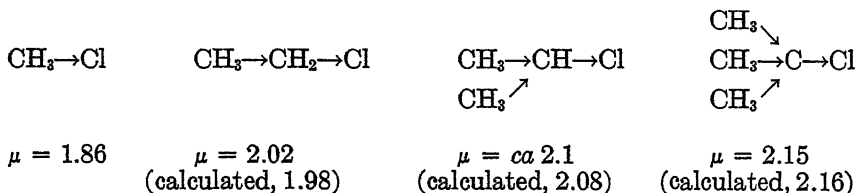


The figures suggest that the dipole moment does indeed vary in the anticipated manner with the position of the substituting atom in the periodic system: the moment increases as the group number increases for elements of the same period, and, on the whole, as the period number decreases for elements of the same group. Two observations, however, are necessary concerning this conclusion. The first is that the absence of an axis of symmetry in the amino and hydroxyl groups must confer on moments of the methylamine and methyl alcohol molecules a component ( $\downarrow$ ) perpendicular to the direction of the linking of the substituent. This perpendicular component will necessarily contribute to (and so increase) the observed resultant moment. That may be why the difference between the moments of ethane and methylamine is so much greater than the difference between the moments of methylamine, methyl alcohol, and methyl fluoride, for obviously there can be no perpendicular components in the moments of ethane and the methyl halides. The second point is that, owing to the different sizes of the halogen atoms, electron displacements equivalent to the same electrical charge transferred from carbon to halogen would lead to dipole moments increasing from methyl fluoride to methyl iodide; wherefore the electrical transfers must actually decrease in this direction more rapidly than do the observed dipole moments.

The influence of multiple linkings may be illustrated by the following comparisons:

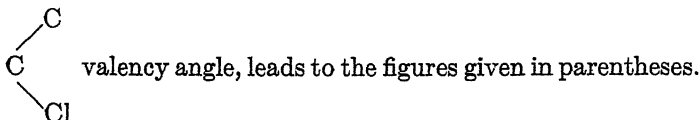


The hypothesis of the propagation of the inductive effect through saturated hydrocarbon chains, or, what amounts to the same thing, the assumption that alkyl groups derive polarity from the presence of intrinsically polar substituents, also finds confirmation in measurements of dipole moments. An illustration is available in the dipole moments of the simple alkyl chlorides:

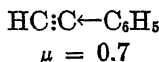
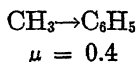
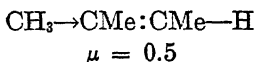


It will be seen that the dipole moments of these substances exhibit a uniform increase with homology. The magnitude of the increase can be accounted for by the assumption that the polarity conferred on each

$\beta$ -methyl group by the  $\alpha$ -chloro substituent is equivalent to a group dipole moment of about 0.3 unit : the adoption of this value, and of  $110^\circ$  for the



The dipole moments associated with the attachment to alkyl groups of ethenyl and phenyl radicals may be illustrated by reference to the determined values for amylene and toluene. The stronger electron attraction of the acetylenyl group is shown by the moment of phenylacetylene. The directions of the moments of toluene and phenylacetylene have been experimentally proved:

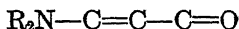


### C. Genesis of the mesomeric effect

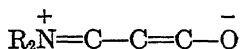
In the theory of the mesomeric effect the curved arrow signs in a formula such as



signify that certain of the electrons are "strained" away from the positions which they would be expected in the absence of such deformations to occupy in the system



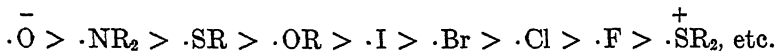
and towards those positions in which the same electrons appear in the alternative structure



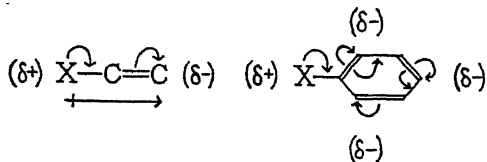
The extent to which this partial transformation is likely to proceed in a given case will evidently depend on the "stability" of the second structure relatively to that of the first. From this point of view, the main difference between the two structures is that one of the terminal atoms has increased its covalency, becoming cationic, whilst the other terminal atom has decreased its covalency, becoming anionic, on passing from the first form to the second. It is therefore necessary to consider the factors which confer on substituents a tendency to increase or decrease their covalency.

The theoretical argument concerning this question (11, 16) is closely similar to that already developed in connection with the inductive effect. Increases of covalency will be first considered, and, since groups with complete valency electron shells can enhance their covalency only if they

possess unshared electrons, the most important of the groups to be considered are those the characteristic atom of which possesses one or more unshared electron pairs. There will be an important distinction between charged and formally neutral groups: negatively charged groups will, on the whole, tend to increase their covalency, becoming neutral, to a greater extent than will neutral groups, which on increasing their covalency must become positive; and neutral groups as a class will, in turn, tend to augment their covalency more than will positively charged groups (e.g.,  $\bar{\text{O}} > \cdot\text{OR} > \cdot\text{OR}_2$ ). Amongst groups in any one state of electrification, there will be a connection with the periodic table. Thus the increasing centralization of nuclear charges in the series  $\cdot\text{NH}_2$ ,  $\cdot\text{OH}$ ,  $\cdot\text{F}$ , implies an increasing nuclear hold on the unshared electrons, and therefore a diminishing tendency to increase covalency ( $\cdot\text{NR}_2 > \cdot\text{OR} > \cdot\text{F}$ ). Again, in the halogen series, the iodine atom being the most deformable should have the greatest tendency to enhance its covalency ( $\cdot\text{I} > \cdot\text{Br} > \cdot\text{Cl} > \cdot\text{F}$ ). These distinctions clearly correspond to known differences of basic character, and it is possible on this empirical basis to construct a series which embodies the comparisons illustrated:



When any of these groups is attached to an unsaturated residue, the tautomeric mechanism of electron displacement will operate always in such a way as to decrease the electron content of the group X, and increase that of the unsaturated residue:

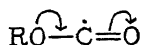


(+M effect)

An effect in this direction is represented by the symbol  $+T$ , and, in so far as the electron displacement has permanent character (mesomeric effect), by the symbol  $+M$ . The mesomeric effect thus corresponds to a molecular dipole, the direction of which is shown by the symbol  $\leftrightarrow$  under the above formulas. The stable positions of the electrons are expected to be further removed from those which they would occupy in the absence of such an effect, the greater the tendency of the system to undergo covalency change. If the groups which exert a  $+M$  effect are imagined to be sever-



it will undoubtedly be definite and considerably greater in, for example,



The group  $\cdot\text{CO}_2\text{R}$  is a composite group of the general form  $\text{X}-\dot{\text{C}}=\text{Y}$ , and it will be useful to consider the way in which the mesomeric effect of such groups is expected to change when X and Y are independently varied.

The effect of the variation of X may be illustrated by reference to the series  $\text{X}-\dot{\text{C}}=\text{O}$ . It will be evident from the foregoing discussion that the strength of the internal effect in groups of this form will depend on the  $+M$  effect of X, and will correspond to such a sequence as the following:

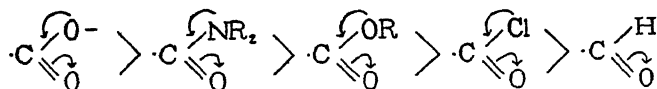
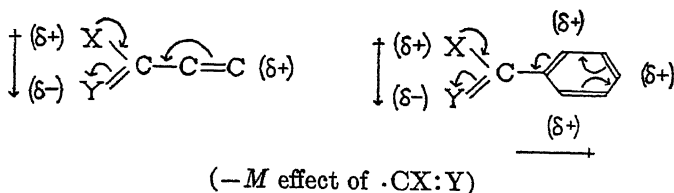


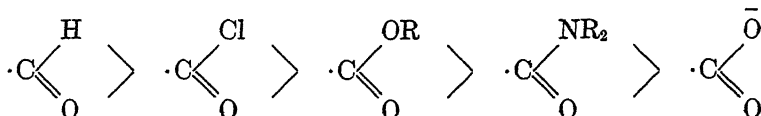
TABLE 2  
*Mesomeric effect*

$+M$ effect of X (electron repulsion dependent on covalency increase)	
$\cdot\text{CR}_2 > \cdot\text{NR} > \cdot\text{O};$	
$\cdot\text{NR} > \cdot\text{NR}_2; \cdot\text{O} > \text{OR} > \text{OR}_2; \cdot\text{S} > \cdot\text{SR} > \text{SR}_2; \cdot\text{I} > \cdot\text{IR};$	
$\cdot\text{NR}_2 > \text{OR} > \cdot\text{F};$	
$\cdot\text{SR} > \cdot\text{OR}; \cdot\text{I} > \cdot\text{Br} > \cdot\text{Cl} > \cdot\text{F}$	
$-M$ effect of Y (electron attraction dependent on covalency decrease)	
$\cdot\text{NR}_2 > \cdot\text{NR};$	
$\cdot\text{O} > \cdot\text{NR} > \cdot\text{CR}_2$ (e.g., $\cdot\text{C}:\text{O} > \cdot\text{C}:\text{NR}; \cdot\text{N}:\text{O} > \cdot\text{N}:\text{NR}$ );	
$\cdot\text{S} > \cdot\text{O};$	
$\cdot\text{N}$ (e.g., in $\cdot\text{C}:\text{N}, \cdot\text{N}:\text{N}$ )	

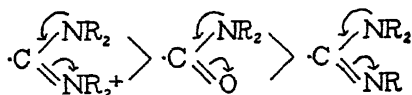
In the last member of this sequence the effect of X (= H) is zero by definition; in the first it is equivalent to the transference of half an electron, because the mesomeric form of the carboxylate ion must obviously be symmetrical with respect to the oxygen atoms. When any of these groups is attached through its carbon atom to an unsaturated system, the mesomeric effect of the whole substituent will be in the  $-M$  direction for evident constitutional reasons:



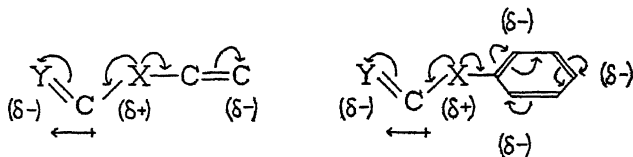
It will be seen from the formulas that the effect of the whole group,  $\cdot CX:Y$ , and the effect within the group are in competition with each other. Therefore the series representing the relative intensities of the  $-M$  effect due to the whole substituent is to be obtained by inverting the series given above:



The complementary comparison, in which X is kept constant whilst Y is varied, may be illustrated by reference to the series  $\text{R}_2\text{N}-\dot{\text{C}}=\text{Y}$ . The relative intensities of the internal effect in groups of this form will be determined by the  $-M$  effect of Y, and may be exemplified by the following typical series:

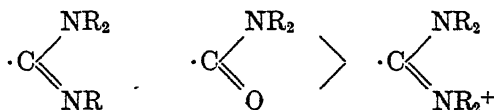


If these groups are attached through their carbon atoms to an unsaturated system, the relative values of the mesomeric effect ( $-M$ ) of the whole substituent will be given by this sequence, since the internal, competing effect of X ( $= \cdot\text{NR}_2$ ) is constant. If, however, the groups are attached to the unsaturated residue, not through carbon, but through nitrogen (in general, through the atom X), the mesomeric effect of the whole group will necessarily be in the  $+M$  direction, as the following formulas show:



( $+M$  effect of  $\cdot\text{XC}:Y$ )

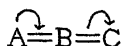
This  $+M$  effect is evidently in competition with the internal effect. It follows that the sequence expressing the relative intensities of the effect of the whole group is to be obtained by inverting the series for the internal effect:





The  $\cdot\text{NO}_2$  group is a composite group of the  $-M$  series; it is closely analogous to the  $\cdot\text{CO}_2$  group, and, like the latter, has a symmetrical mesomeric form. The  $\cdot\text{S}\cdot\text{C}:\text{N}$  group belongs to the  $+M$  series, and may be compared with the groups  $\cdot\text{O}\cdot\text{CR}:\text{O}$  and  $\cdot\text{NR}\cdot\text{CR}:\text{O}$ .

A number of groups of the general form  $\text{A}:\text{B}:\text{C}$  contain a system capable of a mesomeric state by reason of the presence of an unshared valency duplet in one of the atoms A or C. The alternative structure must now contain a triple linking,  $\overset{+}{\text{A}}:\text{B}:\bar{\text{C}}$ , but otherwise the case



is not essentially different from that in which there is only one double bond in the original three-atom system. The characteristic groups of the ketenes and aliphatic diazo compounds furnish examples (see table 3). In azides and cyanates unshared duplets are present in both the terminal atoms, so that there is a minor mesomeric polarization in opposition to the major one. In carbodiimides the symmetry of the system renders the counter-polarizations of about equal intensity, so that the net effect is small.

TABLE 3  
*Mesomeric effect in groups of the form A:B:C*

TYPE	POLARIZATIONS	ALTERNATIVE STRUCTURES
Ketenes.....	$\overset{\curvearrowright}{\text{O}}=\text{C}=\overset{\curvearrowright}{\text{CR}}_2 \quad \dots$	$\overset{+}{\text{O}}\equiv\text{C}-\bar{\text{CR}}_2 \quad \dots$
Diazo compounds.	$\overset{\curvearrowright}{\text{N}}=\overset{+}{\text{N}}=\overset{\curvearrowright}{\text{CR}}_2 \quad \dots$	$\text{N}\equiv\overset{+}{\text{N}}-\bar{\text{CR}}_2 \quad \dots$
Azides ....	$\overset{\curvearrowright}{\text{N}}=\overset{+}{\text{N}}=\overset{\curvearrowright}{\text{NR}} \quad \overset{\curvearrowright}{\text{N}}=\overset{+}{\text{N}}=\overset{\curvearrowright}{\text{NR}}$ (major) (minor)	$\text{N}\equiv\overset{+}{\text{N}}-\bar{\text{NR}} \quad \bar{\text{N}}-\overset{+}{\text{N}}\equiv\overset{+}{\text{NR}}$ (major) (minor)
Cyanates.....	$\overset{\curvearrowright}{\text{O}}=\text{C}=\overset{\curvearrowright}{\text{NR}} \quad \overset{\curvearrowright}{\text{O}}=\text{C}=\overset{\curvearrowright}{\text{NR}}$ (minor) (major)	$\overset{+}{\text{O}}\equiv\text{C}-\bar{\text{NR}} \quad \bar{\text{O}}-\text{C}\equiv\overset{+}{\text{NR}}$ (minor) (major)
Carbodiimides....	$\text{RN}=\overset{\curvearrowright}{\text{C}}=\overset{\curvearrowright}{\text{NR}}' \quad \text{RN}=\overset{\curvearrowright}{\text{C}}=\overset{\curvearrowright}{\text{NR}}'$ (equal)	$\overset{+}{\text{RN}}\equiv\text{C}-\bar{\text{NR}}' \quad \text{RN}-\bar{\text{C}}\equiv\overset{+}{\text{NR}}'$ (equal)

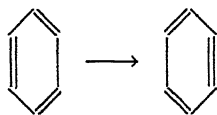
It has been seen that the electrical transference involved in the mesomeric polarization of the carboxylate ion amounts to one-half of that which would be required for conversion into the alternative structure:

$\overset{-}{\text{O}}:\text{C}:\overset{-}{\text{O}} \rightarrow \overset{-}{\text{O}}:\text{C}:\overset{-}{\text{O}}$ . In this transition an ionic center is displaced. In the azide group the symmetry is less perfect, but the mesomeric polarization would still be expected to be of the order of one-half of that which corre-

sponds to a complete interconversion of structures:  $\text{RN}^{\cdot-}\cdot\text{N}^{\cdot+}:\text{N} \rightarrow \text{RN}^{\cdot+}:\text{N}^{\cdot-}:\text{N}$ . In this transition a dipole is displaced. We have seen also that for any formally neutral system of the same general type, for instance the carboxyl-amido group, the mesomeric electrical transference must be definitely less than one-half of that necessary for a complete interchange of structures:

$\text{H}_2\text{N}^{\cdot-}:\text{C}^{\cdot+}:\text{O} \rightarrow \text{H}_2\text{N}^{\cdot+}:\text{C}^{\cdot-}:\text{O}$ . In this transition a dipole is *created*. According to physical evidence to be mentioned later, the usual order of magnitude of the mesomeric polarization of formally neutral systems is about one-tenth of what would correspond to the completed interchange. We may reasonably suppose that in these cases the work of creating the dipole limits the polarization. The smallness of the fraction characterizing the mesomeric polarization of formally neutral systems is important, because it means that the main direction of possible further polarization by the tautomeric mechanism of electron displacement is *towards* the dipolar structure. This is the origin of the strongly unidirectional "valve-like" character of many of the electromeric polarizability effects which are discussed in the sequel.

The limitation imposed by dipole formation on the mesomeric polarizations of the neutral conjugated systems hitherto considered depends on the circumstance that these systems are open; in a closed conjugated system no dipoles would be created by an analogous polarization. In benzene, on account of symmetry, there must be a mesomeric polarization equal to exactly one-half of that required for the interconversion of the alternative Kekulé structures:



The theory of the mesomeric effect thus leads to a benzene hypothesis which may be regarded as a modernized version of both the oscillation formula of Kekulé and the partial valency formula of Thiele; each displaceable electron duplet is shared just as much by  $\text{C}_1\text{C}_2$  as by  $\text{C}_2\text{C}_3$ , and the molecule thus contains a sixfold axis of symmetry. It should be emphasized that in the polysubstitution products of benzene the fraction characterizing the mesomeric polarization of the aromatic nucleus will not in general be exactly one-half; the ring structure will usually approximate more closely to one Kekulé structure than the other, and will thus possess a lower effective symmetry than benzene itself. The same applies in the case of polynuclear aromatic systems.

The only mesomeric polarizations which can be displayed by the unsatu-

rated system present in ethylene and polyenes are those which may arise through interaction between the system and its substituents. Ethylene can be regarded as having two activated structures,  $\overset{+}{\text{C}}=\overset{-}{\text{C}}$  and  $\overset{-}{\text{C}}=\overset{+}{\text{C}}$ , and, in a formal sense, these may separately be correlated with hypothetical mesomeric polarizations which, in any symmetrically substituted ethylene derivative, would mutually cancel. On the other hand, in the presence of a substituent calculated to render one of the activated structures relatively more stable than the other, the condition of the system would begin to approach that of the carbonyl group, and there could be a residual mesomeric polarization:

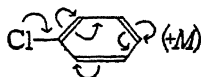


This result is of course the same as that which would have followed directly from the hypothesis that ethenoid linkings readily take up an impressed polarity; but it seems advantageous to connect the facile deformation of unsaturated systems with the general theory of the mesomeric effect.

The polarization of ethenoid systems by attached substituents will be strongly limited by resistance to disturbance of the octets. This limitation will always be important unless the ethenoid linking is incorporated in a more extended conjugated system, when mesomeric polarization may be able to occur without disruption of any octet. Thus in  $\text{Cl}-\text{C}=\text{C}$  there will be two impressed polarizations of the double linking, but that formulated to the left will be the more important because it preserves the octets:



Similarly the dominating mesomeric polarization of  $\text{Cl} \cdot \text{C}_6\text{H}_5$  is



and not



On the other hand, if the ethenoid or polyenoid system contains a second substituent which obviates the octet disruption associated with the subsidiary type of mesomeric polarization, then the relative importance of this effect will be increased, or, in other words, the net  $+M$  effect, due to the chlorine atom, will suffer diminution. Thus in *p*-chloroanisole the  $+M$  polarization of chlorine is diminished by a larger subsidiary  $-M$  polarization than would have been the case had the methoxyl group been absent. That is to say, in *p*-chloroanisole the total mesomeric polarization will not

be equal to the vector sum of the mesomeric polarizations in chlorobenzene and in anisole; there will be a departure from additivity, owing to the mutual interaction of the substituents. This shows that not only the inductive effect but also the mesomeric effect may participate in the general phenomenon of mutual interference between group polarizations.

*D. Source of the energy of mesomerism (valency degeneracy)*

It will have been noticed already that although the theory of octet stability is used in the discussion both of the inductive and of the mesomeric effect, it is applied in two quite different ways; in the former case it is combined with the principle of induction, and in the latter with a different principle, which it remains to consider more fully. The inductive and mesomeric effects have been treated as essentially independent phenomena, and, in anticipation of conclusions reached later, it may be stated that this conception of the independence of the effects is fully borne out by both the physical and the chemical evidence. The implications of this independence can be exhibited by reference to the category of groups (e.g.,  $\cdot\text{OMe}$ ,  $\cdot\text{Cl}$ ) which, when attached to an unsaturated carbon system, exert the complex polar effect summarized in the formula  $-I+M$ : they simultaneously produce oppositely directed polarizations, attracting electrons by one mechanism and repelling them by another, so that the total polarization of the system is really the resultant of these two effects (see tables 1 and 2):



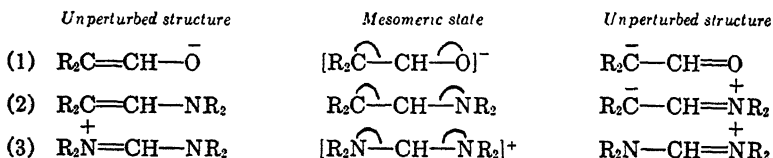
[ $(-I+M)$  effect;  $-I$  represented by  $\curvearrowleft$ , and  $+M$  by curved arrow]

Since the two effects may thus work in opposition, it follows that one must possess some specific source of energy which is denied to the other, and thus the principle underlying the mesomeric effect, which still remains to be explained, is evidently a principle (15) which will provide that effect with an exclusive source of energy.<sup>3</sup>

The question of the nature of this peculiar source of energy may be approached by considering the conditions obtaining in any simple structure capable of the mesomeric effect; for the sake of generality, we may take three examples, namely, an anion, a natural molecule, and a cation. Each entity must have at least two ordinary modes of representation, as is illustrated in the first and third columns of the subjoined scheme of formulas; and the circumstance that some of these structures are charged as a whole whilst others are neutral as a whole, and that transitions between

<sup>3</sup> This conclusion was independently reached by Professor G. E. K. Branch, and has been used by him for several years past in teaching and in colloquia.

the alternative structures may involve a displacement, a neutralization, or even a mutual separation, of local atomic charges is trivial in comparison with the essential similarity of the three relationships exhibited:



From the present point of view the important feature of the analogy is that all three systems show that type of relationship which is expected to lead to a partial or complete quantum mechanical degeneracy. In the first place, each of the pairs of formulas shown in the first and third columns represents the same entity with different electronic arrangements; these structures may be called the "unperturbed" structures. Secondly, in one of these cases the two unperturbed structures have the same energy, and in the others they may be assumed to have sufficiently nearly the same energy to preserve the qualitative features of the analogy. These conditions define a degeneracy (which will be complete if the energies of the unperturbed structures are equal, and incomplete if they are not exactly equal) and, unless there exists some unknown principle "forbidding" transition between the alternative structures, the "perturbation" necessary to produce the actual system will involve separation of electronic energy levels. Some of these new energies will correspond to excited states. The normal, and most stable, state will, however, be one which, on account of the perturbation energy, has less energy than could be ascribed to either of the unperturbed structures. This is the mesomeric state. (It will be appreciated that we are here picturing the production of real states from unreal states, and not the deformation of real states by some external disturbance, as is the case in most of the physical problems to which perturbation theory is commonly applied.) Where there is an exact equivalence of unperturbed energies the mesomeric state must be symmetrically related to the unperturbed structures, whilst in other cases the relations of the mesomeric state to the separate unperturbed structures although not identical will be qualitatively similar (hence the term "mesomeric," from *μέσσω* and *μέρω*). Accordingly, the mesomeric state may be represented as shown in the center column of the above scheme of formulas, in which each symbol  $\curvearrowright$  denotes the distributed wave function of an unlocalized electron-duplet. Thus the specific and exclusive source of energy which is necessary to determine the independence of the mesomeric effect is identified as perturbation energy dependent on a degeneracy (15). The ultimate justification for this hypothesis is, of course, entirely a *poste-*

*riori* in character; it is derived partly from the physical evidence discussed in the sections immediately following, and partly from the necessity for the hypothesis of the mesomeric effect in the interpretation of organic reactions.

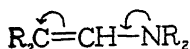
The theory of electron degeneracy in conjugated systems has been developed independently by L. Pauling (34, 35), who has studied in detail from the viewpoint of quantum mechanics the special cases of cyclic conjugation encountered in benzene and other aromatic systems; he has been able, indeed, to evolve a method of treatment which in principle enables the perturbation energies of conjugated unsaturated hydrocarbons to be calculated on a comparative basis. Pauling describes the phenomenon under the name "resonance," which, as is well known, is based on the mathematical analogy between mechanical resonance and the behavior of wave functions in quantum mechanical exchange phenomena. There appears, however, to be some possibility that this method of description may suggest an analogy which has never been intended. It is clearly an essential point in the theory of valency degeneracy that the energy effect by which the system acquires increased stability corresponds to new electron motions in spatially modified wave functions. On the other hand, the idea appears to have gained some ground recently that the conception of a mesomeric state is unnecessary, that the unperturbed structures are all that really exist, and that these pass into each other "like tautomerides but much more rapidly," the great frequency of interchange accounting for the energy effect. Now one of the characteristics of a tautomeric system is the existence of forms which clearly correspond to separate molecular states, because each molecule spends substantially the whole of its life partly in one form and partly in the other, and only a quite insignificant fraction of its life in the actual process of transition. If, however, in any of the most typical cases of "resonance," we attempt to interpret the gain in stability by assuming a tautomerism of unperturbed states, the frequency of interchange which it is necessary to postulate in order to account for the known magnitude of the energy effect may be so great as to require that every molecule must spend most or all of its time in changing, and cannot remain quiescent for significant periods in either of the assumed states; in other words, the term "state" loses its meaning in reference to the only molecular states which this theory recognizes. The assumed frequency of valency interchange is, indeed, of about the same sort of magnitude as the frequencies which are attributable generally to combined electrons independently of valency resonance. There can be no physical separation, therefore, between resonance vibrations and other electronic vibrations; it follows that the unperturbed structures, in which the resonance frequencies are absent by assumption, are only of the nature of intellectual scaffold-

ing, and that the actual state is the mesomeric state. The fact that a condition of constant electronic vibration is emphasized in the description of this state does not differentiate it from, but rather classifies it with, the normal molecular states which occur independently of resonance. The additional electronic energy associated with resonance naturally implies an altered wave function, but the reasons why we do not associate this energy difference with a definite frequency along a definite path are quite analogous to the reasons for not reverting to Bohr orbits in the description of molecular structures generally.

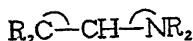
The chemical evidence in support of the above ideas is so extensive that it could be dealt with adequately only in a review of the applications of the theory to various types of organic reactions; some general remarks may, however, be included here. In the first place, it may be noted that the examples in the above scheme of formulas (see p. 250) typify three extensive fields in organic chemistry, and that in each case there is a mass of evidence pointing to the non-existence, as separate individuals, of such isomeric modifications as those represented in the unperturbed structures. In example 1 these structures correspond to the hypothetical anions of a mobile hydrogen tautomeric system, and it cannot be regarded as accidental that in this very fully explored domain numerous pairs of covalently constituted hydrogen isomerides are known with very widely differing velocities of interchange, but no single pair of corresponding alkali metal salts; in no case is there any experimental evidence of the existence of more than one anion. The same general indication emerges from the study of the type of tautomerism represented by example 3, although here the field of observation is not so wide. Concerning example 2 it is a commonplace that isomerides corresponding to the unperturbed structures are unknown amongst neutral molecules. All this is negative evidence; it can still be argued that in every investigated example tautomerism between unperturbed forms takes place too rapidly for experimental detection. But the consideration of any simple problem of reactivity in which mesomerism plays a leading part shows that this explanation is inadequate. For example, aniline is a weaker base by about a millionfold than any primary aliphatic amine such as methylamine or *tertiary*-butylamine. Now the neutral unperturbed structure for aniline,  $\text{NH}_2\text{—C}_6\text{H}_5$ , requires a basicity of about the same order of magnitude as that of an aliphatic amine; and the three dipolar unperturbed structures, which may collectively be represented  $\text{NH}_2^+\text{=C}_6\text{H}_5^-$ , all require zero basicity for ammonium salt formation. If we were to try to account for the actual basicity of aniline by postulating a *tautomerism*, too rapid for direct detection, between these unperturbed structures considered as molecular states, we should be led to the ludicrous

conclusion that the molecule exists almost entirely in the betaine forms. The only way to avoid this difficulty would be to increase the assumed rate of interchange to such a degree that the molecules would usually fail to remain in the most basic form,  $\text{NH}_2\text{—C}_6\text{H}_5$ , for the duration of a molecular collision; for if this were so, even a high instantaneous concentration of  $\text{NH}_2\text{—C}_6\text{H}_5$  molecules would be unable to produce a corresponding amount of basic reactivity. But this supposition is tantamount to discarding altogether the idea of unperturbed forms as molecular states, and adopting in its place the conception of a state distinct in properties from either of the states originally assumed.

A remark may be added concerning notation. There are two ways in which a mesomeric state may be symbolized. In the first place, one of the unperturbed structures may be written down, and then curved arrow symbols may be attached to it for the purpose of indicating how it has to be "corrected" in order to represent the actual state of the system, e.g.,



(If this notation is used it must be made clear otherwise that a mesomeric polarization and not an electromeric polarizability effect is intended, for the curved arrows themselves denote only a *mechanism* of electron displacement, independently of whether this is supposed to characterize a molecular state or a process occurring in reaction.) Alternatively the "scaffolding" may be discarded altogether and a curved bond sign may be introduced for each distributed duplet, e.g.,



It is convenient to be able to employ either alternative according to circumstances.

#### E. Mesomerism and the evidence of dipole moment measurements

The use of dipole moment measurements as a physical test for the mesomeric effect was first proposed (12) in 1926 in advance of the appearance of appropriate data. The test depended on finding a group ( $\cdot\text{NMe}_2$  was suggested) such that when it is linked to an unsaturated system, e.g., a phenyl group, the mesomeric effect opposes, and is strong enough to outweigh, the inductive effect; it then had to be shown by the use of the vector addition principle that the direction of the electric moment associated with the saturated aliphatic linking of the group becomes reversed in the corresponding aromatic combination:





The first test of this kind was completed (for the group  $\cdot\text{NH}_2$ ) in 1928 by K. Höjendahl (10) with definitely positive results, subsequently confirmed by other workers.

Extensive further confirmation has emerged from a more general form of the test which was applied in 1931 by L. E. Sutton (41). He showed that, even when (as is general) the mesomeric effect does not actually outweigh the inductive, there are highly significant differences between the moments associated with the types Alkyl—R and Aryl—R: the direction of these differences always corresponds to the direction of the mesomeric effect in the system Aryl—R. Comparing, as far as the data allowed, *tertiary*-alkyl with phenyl compounds, Sutton obtained the differences noted in table 4, in calculating which molecular dipoles with their positive ends in the group R are taken as positive, and those in the opposite direction negative, the subtraction being algebraic.<sup>4</sup>

TABLE 4  
*Differences in moments of aliphatic and aromatic derivatives*  
 $\pm$  signs correspond to  $\pm M$  effects

GROUP R	$\mu (\text{Ar}\cdot\text{R}) - \mu (\text{Alk}\cdot\text{R})$	GROUP R	$\mu (\text{Ar}\cdot\text{R}) - \mu (\text{Alk}\cdot\text{R})$	GROUP R	$\mu (\text{Ar}\cdot\text{R}) - \mu (\text{Alk}\cdot\text{R})$
$\cdot\text{NH}_2$	+ (0.32)	$\cdot\text{Br}$	+0.69	$\cdot\text{CO}$	— (0.28)
$\cdot\text{S}\cdot$	+ (0.10)	$\cdot\text{Cl}$	+0.59	$\cdot\text{CHO}$	— (0.29)
$\cdot\text{O}\cdot$	+ (0.23)	$\cdot\text{CH}_3$	+0.45	$\cdot\text{COMe}$	— (0.18)
$\cdot\text{OMe}$	+ (0.0)	$\cdot\text{CH}_2\text{Cl}$	+ (0.21)	$\cdot\text{CO}_2\text{Me}$	— (0.22)
$\cdot\text{OH}$	+ (0.15)	$\cdot\text{CHCl}_2$	$\pm$ (0.0)	$\cdot\text{CN}$	—0.43
$\cdot\text{I}$	+0.86	$\cdot\text{CCl}_3$	—0.50	$\cdot\text{NO}_2$	—0.88

The reason why certain of the numerical values are placed in parentheses in this table is as follows. In a strict application of the principles underlying the test, the subtraction of molecular dipole moments should be, as Sutton recognized, not an algebraic, but a vectorial, subtraction. But when the bonding direction of R is an axis of symmetry, the molecular dipole moment must lie in the direction of this axis, and algebraic and vectorial subtractions give identical results, so that, for the groups  $\cdot\text{I}$ ,  $\cdot\text{Br}$ ,  $\cdot\text{Cl}$ ,  $\cdot\text{CH}_3$ ,  $\cdot\text{CCl}_3$ ,  $\cdot\text{CN}$ ,  $\cdot\text{NO}_2$  (the experimental reason for regarding the nitro group as axially symmetrical is given below), the recorded differences may be regarded as true measures of the mesomeric effect, correct

<sup>4</sup> L. E. Sutton noted the correspondence of the signs of his differences to the directions assumed for electromeric activations in the aromatic compounds, but overlooked the circumstance that *permanent* electron displacements by the tautomeric mechanism in the same directions had previously been postulated by C. K. Ingold and E. H. Ingold (1926) as a general property of unsaturated systems.

both as to sign and magnitude within the limits of experimental precision. On the other hand, when the bonding direction of R is not an axis of symmetry, then the molecular dipole axes must lie in other directions, which at present are not accurately known. Thus a strict vectorial subtraction is impossible, and all that can be said concerning the results of the algebraic subtraction by which it is replaced in the above table is that on the whole the signs are still important, although the numerical values are without significance. It should be added that even the signs of the differences for unsymmetrical groups cannot be guaranteed. For instance, they could easily lose their meaning if the induction dipoles had widely different directions in corresponding aliphatic and aromatic compounds; this is what would happen if the valency angle of the substituent atom changed considerably from one type of compound to the other. Valency angles are known to change with alterations in the groups which the valencies bind, but the variations are usually rather small, and it may be assumed that in the present cases they are insufficient to cause serious disturbance to the signs of the dipole differences.

Until we have more accurate information about valency angles, it is the data for the axially symmetrical groups which are of chief interest:

$\cdot\text{I}$	$\cdot\text{Br}$	$\cdot\text{Cl}$	$\cdot\text{CH}_3$	$\cdot\text{CCl}_3$	$\cdot\text{CN}$	$\cdot\text{NO}_2$
+0.86	+0.69	+0.59	+0.45	-0.50	-0.43	-0.88

These figures clearly show that the mesomeric dipole moments are small in comparison with the moments which would be possessed by the corresponding dipolar structures (of the order of 10 units). It is also to be observed that the numerical values give a correct sequence, and, in particular, that the halogens, although electron-attracting, have a  $+M$  effect. Again, the groups  $\cdot\text{Cl}$  and  $\cdot\text{CCl}_3$ , although both are electron-attracting, fall into different categories, one having a  $+M$  and the other a  $-M$  effect, whilst electron-repelling  $\cdot\text{CH}_3$  has a mesomeric effect of the same sign as that of electron-attracting  $\cdot\text{Cl}$ . All these relationships are normal and had been theoretically anticipated. It is of interest that the expected interaction between the mesomeric effects of two substituents in the same molecule, illustrated in the preceding section by an analysis of the example of *p*-chloroanisole, has recently been established from dipole moment data by G. M. Bennett and S. Glasstone (2).

It is important to notice that these observations go a long way towards confirming the conclusion that mesomerism cannot be interpreted as a tautomerism between molecular states corresponding to the unperturbed structures. It is known that measured dipole moments undergo no change with increasing frequency of the alternating field used for the measurement until this frequency corresponds in order of magnitude to the minimum

time (measured by the so-called "relaxation time") required by the polar molecule to reorient itself with each reversal of the field. This time is of the order of  $10^{-12}$  second. Dipole measurements may therefore be regarded as dependent on a physical reaction (the orientation of the molecule by the electric field), the duration of which is of this order of magnitude. Tautomerides of longer average life than  $10^{-12}$  second should independently be able to follow the reversals of the measuring field, and should therefore contribute to any observed moment in accordance with their concentrations and the numerical values of their individual moments, the directions of these moments being immaterial. The dipole moment of any tautomeric mixture should therefore always be numerically between (as distinct from vectorially between) the separate moments of the tautomerides. Now it is not by any means invariably the case that the dipole moments of mesomeric compounds are numerically bracketed between the moments which should be possessed by their unperturbed structures. Thus iodobenzene has one formally neutral unperturbed structure,  $\text{I}-\text{C}_6\text{H}_5$ , the numerical magnitude of the moment of which should, according to Sutton, be 2.13

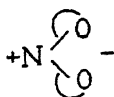
units; it also has three betaïne-like structures of the type  $\overset{+}{\text{I}}=\overset{-}{\text{C}_6\text{H}_5}$ , the moments of which would each be of the order of 10 units. If these structures represented tautomerides, the numerical value of the observed moment of iodobenzene should lie between 2.13 and about 10 units, that is, it should be above 2.13 units; actually it is only 1.27 units. Similar remarks apply to bromobenzene and chlorobenzene. This shows that when we try to give to the unperturbed structures the status of tautomerides, we are forced to assume that their average lives must be less than  $10^{-12}$  second. But this is almost of the order of the duration of a molecular collision, and we have already seen that tautomerides with lives of such a length could not be regarded as representing distinct molecular states. The corresponding requirement of the theory of the mesomeric state is, of course, that the observed dipole value should lie *vectorially* between the values corresponding to the unperturbed structures. This is always found to be the case where the directions of the moments are known; thus, for iodobenzene,  $-1.27$  lies between  $-2.13$  and  $+10$ .

The same point may be illustrated by some examples of a different type, which have been studied by N. V. Sidgwick and his collaborators (39). In aliphatic diazo compounds, for instance, the unperturbed structures,  $\overset{+}{\text{CR}_2}:\overset{-}{\text{N}}:\overset{-}{\text{N}}$  and  $\overset{-}{\text{CR}_2}:\overset{+}{\text{N}}:\overset{+}{\text{N}}$ , each contain a formal heteropolar system, and each should therefore correspond to a moment of about 4 units. The fact that these formal dipoles have opposite directions should not affect the situation if the unperturbed structures are tautomerides; the observed moment should be a numerical mean, namely, about 4 units. Actually,

the observed value is only 1.42 units, and this is clearly to be regarded as a vectorial mean between large, but oppositely directed, unperturbed moments. Similarly in the case of azides, the principal unperturbed structures,  $\text{NR}^+\text{:}\overset{-}{\text{N}}\text{:}\overset{-}{\text{N}}$  and  $\overset{-}{\text{N}}\text{:}\overset{-}{\text{N}}\text{:}\overset{+}{\text{N}}\text{R}$ , possess large oppositely directed dipoles, each of magnitude about 4 units; but the observed moment has the small value 1.55 units. The conclusion to be drawn is exactly the same as that derived from the aromatic examples.

According to the ordinary conception of the nitro group,  $\cdot\overset{+}{\text{N}}\begin{array}{c} \text{O}^- \\ \text{O} \end{array}$ ,

*p*-dinitrobenzene and 4,4'-dinitrodiphenyl should each possess a large dipole moment directed perpendicularly to the line of the nitroxyl groups. This follows from the principle of the rotation of groups about a single linking; for of the infinitude of possible relative rotatory orientations of the nitroxyl groups, all but one must confer on the molecule a finite electric moment. This principle has been fully established by the observed finite dipole moments of many *p*-disubstitution products of benzene with identical substituents the bonding directions of which are not axes of symmetry. It has, however, been found by J. W. Williams (44) that the dipole moments of *p*-dinitrobenzene and of 4,4'-dinitrodiphenyl are indistinguishable from zero, and the same point has been proved for 1,3,5-trinitrobenzene by A. Parts (31) and H. Lütgert (29). It follows that the bonding direction of the nitro group is an axis of symmetry for the group, and therefore that this group is symmetrical with respect to its oxygen atoms as the general theory of the mesomeric effect requires,



#### *F. Mesomerism and the evidence of infra-red spectroscopy*

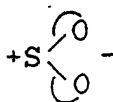
It is a commonplace that the study of infra-red absorption spectra can yield important information concerning certain aspects of molecular structure. The frequencies of the vibration bands measure the forces which control atomic vibrations, whilst the frequencies of the rotation bands and of the rotational structure associated with the vibration bands determine the molecular moments of inertia. In many cases definite conclusions can be drawn concerning the conformation and symmetry of the molecule. Now we have seen that in certain systems, such as the nitro group and the carboxylate ion, mesomerism should impart to the structure an element of symmetry which would not have been expected from older views of mo-

lecular constitution. Such additional symmetry should, if present, be demonstrable by the methods of infra-red spectroscopy, provided the system is one to which those methods can be applied. In practice, application is restricted to molecules containing only a very few atoms, owing to the difficulty of analyzing the highly complex spectra of polyatomic molecules; nevertheless, a small number of simple mesomeric systems exist to which the application of the spectroscopic test is possible.

The simplest structures capable of mesomerism are of necessity triatomic, and the condition that mesomerism, if present in a triatomic system, should produce additional symmetry limits the choice of examples to molecules of the form  $X \cdot Y : X$ , where  $X$  contains unshared electrons. One of the most obvious examples is sulfur dioxide, the unperturbed structure for which,

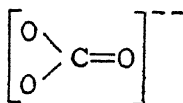


has no plane of symmetry which does not pass through all three atoms. The mesomeric form,

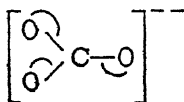


on the other hand, has a plane of symmetry perpendicular to the line of the oxygen atoms. C. R. Bailey and A. B. D. Cassie (1) have shown from an analysis of the infra-red spectrum of sulfur dioxide that the molecule does indeed possess this symmetry, and that the vibrations of the two oxygen atoms are controlled by identical force constants.

A slightly more complicated case is provided by the tetratomic carbonate ion. The unperturbed structure,



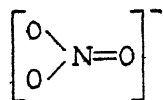
has various planes of symmetry together with a twofold axis. On the other hand, the mesomeric form,



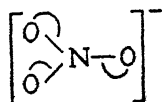
has all these elements of symmetry, and it possesses in addition a threefold axis. It has long been known, from the results of x-ray investigations,

that the carbonate ion has geometrically trigonal symmetry in the crystal. C. Schaefer (37) has shown, however, from a study of the infra-red spectrum, that the ion is also mechanically trigonal in character, the three oxygen atoms being linked by bonds with identical force constants. Moreover, one of the four fundamental frequencies was "optically inactive," indicating a normal vibration with an invariant dipole. The only possible interpretation consistent with the other evidence is that this normal vibration is composed of three symmetrically directed, mechanically and electrically equal, vibrations occurring in phase; and the conclusion follows that the ion is trigonally symmetrical not only with respect to its geometrical and mechanical characteristics, but also with respect to its electrical distribution,—the three oxygen atoms having identical charges and the links identical dipoles.

Schaefer has also investigated the nitrate ion, which is an exactly similar case. The unperturbed structure is



and the mesomeric form,



possesses trigonal symmetry. The ion has four fundamental frequencies, of which one is optically inactive, proving the complete equivalence of the three oxygen atoms (37).

It may be pointed out that such observations as these cannot be interpreted by assuming a tautomerism between molecular states corresponding to the unperturbed structures. For if such tautomerides existed with average lives longer than the period of the atomic vibrations (which is of the order of  $10^{-13.5}$  second), then the vibrations would be characterized by the force constants of the single and double links present in the unperturbed structures, and the symmetry which is characteristic of the mesomeric state would not be found. Thus, any attempt to treat the unperturbed structures as tautomerides would lead to the conclusion that their average lives must be shorter than  $10^{-13.5}$  second, an inference which is inconsistent with the original assumption; for it has already been noted that tautomerides with lives of such a length could not be considered to represent distinct molecular states.

*G. Mesomerism and the evidence of thermochemistry*

The determination of heats of formation by the combination of heats of combustion with the heats of formation of the combustion products from their elements and the heats of atomization of the latter has recently become possible, mainly by reason of the extension of our knowledge of

TABLE 5  
*Perturbation energies*

(i) Carboxyl series		(ii) Aryl series					
SYSTEM	P.E.	SUBSTANCE	P.E.	SUBSTANCE	P.E.	SUBSTANCE	P.E.
·CO <sub>2</sub> R.....	1.2*	Benzene.....	1.7*	Pyridine....	1.9	Pyrrole.....	1.0
·CO·NH <sub>2</sub> .....	1.0*	Naphthalene..	3.2*	Quinoline...	3.0	Indole.....	2.3
CO(OR) <sub>2</sub> .....	1.8	Anthracene...	4.5	—	—	Carbazole...	3.9
CO(NH <sub>2</sub> ) <sub>2</sub> ....	1.6	Phenanthrene..	4.8	Q.M. CALCULATED			
(R = H or alkyl)		Chrysene.....	6.5	Benzene....	1.67	Naphthalene..	3.16

(iii) Aryl-ethenoid and aryl-aryl series				
SUBSTANCE	P.E.	ALLOWANCE FOR SYSTEMS	EXTRA P.E.	Q.M. CALCULATED
Styrene.....	2.0	Ph = 1.7	0.3	0.30
as-Diphenylethylene.....	3.9	2 Ph = 3.4	0.5	0.53
Stilbene.....	4.1	2 Ph = 3.4	0.7	0.58
Diphenyl.....	3.8	2 Ph = 3.4	0.4	0.27

(iv) Substituted aryl series			
SUBSTANCE	P.E.	ALLOWANCE FOR SYSTEMS	EXTRA P.E.
Ph·OR.....	2.0*	Ph = 1.7	0.3
Ph·NR <sub>2</sub> .....	1.9*	Ph = 1.7	0.2
C <sub>6</sub> H <sub>4</sub> (OR) <sub>2</sub> .....	2.5	Ph = 1.7	0.8
Ph·NR·Ph.....	4.0	2 Ph = 3.4	0.6
Ph·COR.....	1.9	Ph = 1.7	0.2
Ph·CN.....	1.9	Ph = 1.7	0.2
Ph·CO <sub>2</sub> R.....	3.1	Ph = 1.7; CO <sub>2</sub> R = 1.2	0.2
Ph·CO·Ph.....	3.9	2 Ph = 3.4	0.5

\* Mean from several derivatives.

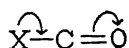
heats of atomization which has accrued from the study of band spectra. Examination of the heats of formation thus determined has shown that a large number of these may be represented, with an accuracy equal to that of the thermochemical measurements, by first ascribing to each type of bond (e.g., C—C, C=C, C—O, etc.) a characteristic heat of formation or

"bond energy," and then summing these bond energies for all the bonds in the molecule.

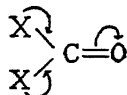
L. Pauling and J. Sherman (33) pointed out that this additive relationship ought not to hold true for valency-degenerate systems, for the heats of formation of these should include their perturbation energies, and should therefore be greater than the sum of the heats of formation of the separate bonds. Collecting together the heats of formation of a large number of substances, these authors observed that whilst many obeyed the additive principle many others did not, and the compounds which did not were just those in which mesomerism was to be expected for constitutional reasons. Moreover the direction of the deviations was always the same, namely, that which corresponds to an increase of molecular stability. It was inferred that the differences were to be interpreted as perturbation energies.

The numerical values of these thermochemically determined perturbation energies are very interesting; some of them are quoted in table 5. The units are electron-volts (e.v.) per molecule, and the figures are accurate to about  $\pm 0.1$  of these units (1 electron-volt per molecule is equivalent to 23.054 kilogram-calories per gram-molecule).

Section i of table 5 shows perturbation energies (P.E.) obtained for various carboxyloid systems. It will be observed that the energy for the perturbation



becomes considerably increased in the more extended perturbation



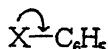
Section ii of table 5 shows the perturbation energies of various aromatic systems. In the homocyclic series the value is 1.7 units for benzene itself, and about 1.5 units for every additional "fused" benzene ring. The degeneracy of pyridine is of the same type as that of benzene, and accordingly the values for pyridine and quinoline are nearly the same as those for benzene and naphthalene, respectively. The degeneracy of the pyrrole ring is of a different type, and it has a lower energy value, which is reflected also in the values obtained for indole and carbazole. The "calculated" values (*italicized*) for benzene and naphthalene are obtained by a method mentioned later.

Section iii of table 5 shows the energy effects of conjugation between an aryl group and an ethylene linking, and between two directly linked aryl groups. In styrene, the thermochemically determined perturbation

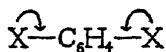


energy (P.E.) is 2.0 units, and this is greater than the value, 1.7, appropriate to the phenyl group; the *extra* perturbation energy (extra P. E.), 0.3 unit, represents the effect of the extension of the purely aromatic degenerate system by the inclusion of perturbations involving the extracyclic double linking. As expected, *as*-diphenylethylene and stilbene show similar but greater effects, and diphenyl shows a smaller one than stilbene. The "calculated" extra perturbation energies, shown in italics in the last column of this section of table 5, are referred to below.

Section iv of table 5 shows the energy effect of mesomeric interaction between the benzene ring and various substituents. In the first four examples, the substituents are of the  $+T$  type, and it will be noticed that the *extra* perturbation energy, due to the process



becomes much increased in those structures in which the effect is duplicated:



or



Similar remarks apply to the last four examples in which the substituents belong to the  $-T$  category.

Sections ii and iii of table 5 contain some italicized figures under the heading "Q.M. calculated"; these are perturbation energies calculated, excepting for a single arbitrary factor, by a quantum mechanical approximation method due to L. Pauling (32). A full description of this method cannot be given here. It will suffice to state that in principle the perturbation energy of any degenerate molecular system can be calculated in terms of certain "exchange" energy integrals. These integrals, however, have not yet been theoretically evaluated, and they must therefore be treated as disposable constants having the dimensions of energy. In principle, the calculation should introduce one exchange integral for each non-equivalent pair of directly linked atoms involved in the degeneracy; but it is considered that the values of these integrals will depend mainly on the nature of the linked atoms, and that, in a first approximation, one and the same integral can be used for all the linked atom-pairs participating in the degeneracy of *hydrocarbons*. Thus the perturbation energies of all conjugated unsaturated hydrocarbons can be calculated with the introduction of but *one* arbitrary constant.

Pauling and Wheland (35) carried out this calculation for benzene and naphthalene and found perturbation energies equal to  $1.1055 \alpha$  and  $2.1053 \alpha$ , respectively, where  $\alpha$  is the unknown C.C-exchange integral. The adoption of a single value for  $\alpha$ , namely 1.5 electron-volts, brings both these energies into agreement with the thermochemically determined values (see table 5, section ii).

Similar calculations have been carried through by Pauling and Sherman (34) for styrene, *as*-diphenylethylene, stilbene, and diphenyl; they obtained for the *extra* perturbation energies of these compounds  $0.201 \alpha$ ,  $0.355 \alpha$ ,  $0.389 \alpha$ , and  $0.117 \alpha$ , respectively. The same value of  $\alpha$ , namely 1.5 electron-volts, brings all these energies into satisfactory agreement with the thermochemical values (see table 5, section iii). Obviously these calculations lend great additional weight to the thermochemical evidence.

TABLE 6  
 $\pm$  Inductomeric polarizabilities

---

$\begin{array}{l} - \\ \cdot O > \cdot OR > \cdot OR_2; \end{array}$	$\begin{array}{l} + \\ \cdot NR_2 > \cdot NR_3; \end{array}$	$\begin{array}{l} + \\ \cdot NR_3; \end{array}$	etc.
$\cdot CR_3 > \cdot NR_2 > \cdot OR > \cdot F;$	etc.		
$\cdot I > \cdot Br > \cdot Cl > \cdot F;$	etc.		
$\cdot CR_3 > \cdot CHR_2 > \cdot CH_2R > \cdot CH_3;$	$\cdot CH_3 > \cdot H;$	etc.	

---

#### H. Inductomeric polarizability

Polarizability effects result from the interaction of two factors, namely the polarizability of the molecular system and the polarizing field of the attacking reagent. The polarizability factor will first be discussed, and the inductomeric and electromeric mechanisms will be considered separately in this connection.

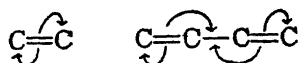
By inductomeric polarizability is understood the polarizability which atoms or groups exhibit in saturated combinations along the lines of their linkings (15). Such polarizability may be expected to decrease with increasing positive electrification of the characteristic atom of the substituent, to decrease with increasing group number, and to increase with increasing period number, in the periodic classification of elements. The argument here is closely similar to one given previously (p. 237), and the relationships thus inferred are summarized in table 6. It is provisionally assumed that an inductomeric polarizability along any linking is the same for both directions; this would surely hold for small electron displacements, but it is unlikely to be more than roughly true for displacements of the magnitude of those which occur during reactions.

Physical data are not of much help for the purpose of checking these conclusions. The atomic polarizabilities derivable from refraction con-

stants are, of course, average values for all directions in space; they also take account of all the electrons, and not only those which constitute bonds. Insufficient is known about molecular anisotropy to permit a calculation of longitudinal bond contributions, and even if these were deduced they would only be limiting values for small displacements. It is nevertheless true that, for elements with closed valency shells, atomic refractions vary over the periodic table in just the manner expected for inductomeric polarizability, with one exception. This relates to an anomalous increase in refraction which is obtained on passing from group IV to group V; the peculiarity may, however, be connected with the fact that it is just at this point that unshared valency electrons first appear.

### I. Electromeric polarizability

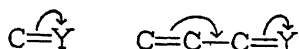
In unsaturated systems electromeric polarizability effects become important, and they have strongly marked directional properties. Unsaturated carbon systems *per se* should be polarizable in either of the two directions indicated:



However, the presence of a substituent of the  $+T$  category renders one of these directions much more important than the other, the whole system deriving properties analogous to those of a non-return valve; for it has already been observed (p. 247) that in neutral mesomeric systems the main direction of possible further polarization is *towards* the formal dipolar structure:



This strongly unidirectional character of electromeric polarizabilities is of considerable importance, because it means that systems such as those illustrated will undergo marked deformation in the field of an attacking reagent only when that field is of the correct polar sign (17). In the presence of substituents of the  $-T$  class the favored direction of electromeric polarizability will, of course, be the opposite:



The magnitude of the electromeric polarizability of an unsaturated system will evidently depend on the energy difference between the mesomeric state and the formal dipolar structure, and thus the effect of groups in modifying the polarizability of an unsaturated carbon system will depend on the factors already considered (p. 242) which militate for or against

covalency change. The principal conclusions concerning these considerations are summarized in table 7 (11, 14, 16, 17, 36).

Once again the evidence of refractivities is fragmentary, mainly owing to the impossibility of analyzing the data at all adequately without more knowledge concerning molecular anisotropy. On the other hand, the fact that electromeric polarizability is of significance in connection with molecular refraction is clearly shown, firstly by the special allowances which are made in order to take account of the effect of simple and conjugated unsaturated systems, and secondly by the variations which the so-called atomic constants exhibit in the presence of such systems. Small variations of this kind are often concealed in the averaging processes whereby these constants are calculated, but large variations cannot thus be obscured, as may be illustrated by a comparison of the "constant"  $[R_L]_D$ , for nitrogen in  $\text{NH}_2 \cdot \text{Alphyl}$  (2.45),  $\text{NH}_2 \cdot \text{Aryl}$  (3.21), and  $\text{NH}_2 \cdot \text{C}:\text{C}:\text{C}:\text{O}$  (4.88).

TABLE 7  
*Electromeric polarizabilities*

---

$+E$ type	
$-$	$+$
$\cdot\text{O} > \cdot\text{OR} > \cdot\text{OR}_2$ ; etc.	
$\cdot\text{NR}_2 > \cdot\text{OR} > \cdot\text{F}$ ; etc.	
$\cdot\text{I} > \cdot\text{Br} > \cdot\text{Cl} > \cdot\text{F}$ ; $\cdot\text{SCN}$ ; etc.	
$-E$ type	
$+$	
$\cdot\text{C}:\text{NR}_2 > \cdot\text{C}:\text{NR}$ ; etc.	
$\cdot\text{C}:\text{O} > \cdot\text{C}:\text{NR}$ ; etc.	
$\cdot\text{COR} > \cdot\text{COCl} > \cdot\text{CO}_2\text{R} > \cdot\text{CO}:\text{NR}_2 > \cdot\text{CO}_2$ ;	$-$
$\cdot\text{CN}$ ; $\cdot\text{NO}_2$ ; etc.	
$\pm E$ type	
$\cdot\text{C}=\text{C}$ ; $\cdot\text{C}=\text{C}-\text{C}=\text{C}$ ; $\text{C}_6\text{H}_5\cdot$ ; etc.	

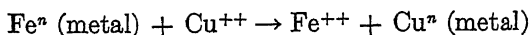
---

## V. POLAR CLASSIFICATION OF REAGENTS

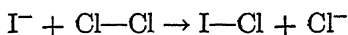
### A. Generalization of the conception of oxidation-reduction

Since it is proposed to regard chemical reactions as electrical transactions in which reagents act by reason of a constitutional affinity either for electrons or for atomic nuclei, it is important to be able to recognize which type of reactivity any given reagent exhibits. For this purpose a correlation must be established between the presumed electronic classification of reagents and their known chemical behavior. The necessary conceptions have been indicated by H. S. Fry (9) and J. Stieglitz (40) especially, and may be developed as follows. An electronic formulation of any oxidation-reduction shows that the reducing agent donates electrons to the oxidizing

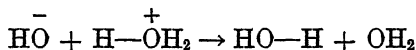
agent. The nature of the exchange is clearest when there is an actual transference of one or more electrons, as in the example:



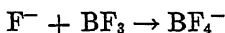
On the other hand, there are many reactions, also recognizably of the oxidation-reduction type, in which the reducing agent does not actually part from the electrons which it donates; it offers only a share in certain of its electrons, e.g.:



The corresponding statement about the oxidizing agent is that it does not in these cases appropriate the electrons which it receives; it acquires merely a share in electrons which previously belonged solely to the reducing agent. In other cases of the donation and reception of a share in an electron-pair, the connection with oxidation-reduction is less direct, but still discernible. Thus in the example:



the hydroxide ion on combination becomes formally neutral; if it were *free* and neutral, it would be a hydroxyl radical, obviously an oxidized form of the original hydroxide ion, which must therefore have acted analogously to a reducing agent. Similarly the proton, derived from its hydrate the oxonium ion, becomes a combined, neutral hydrogen atom; if it were a free, neutral hydrogen atom, it would at once be recognized as a reduced form of the original hydrogen ion, which accordingly must have acted like an oxidizing agent. The reaction:



is clearly of the same general type, although the analogy with oxidation-reduction is still more indirect. These examples adumbrate a series, which by the inclusion of other instances could be rendered more continuous, and the conclusion follows that reduction and oxidation are to be regarded as the prototypes of a more general classification based on the donation and reception of electrons.

### B. Nucleophilic reagents

Reagents which donate their electrons to, or share them with, a foreign atomic nucleus may be termed *nucleophilic* (15); table 8 contains some examples.

The first three reagents of this list act by donating one or more of their electrons completely; they are typical reducing agents in the restricted

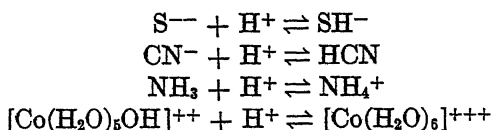
sense of the term. The remaining six examples (Nos. 4 to 9) act by donating a share in one or more of their unshared duplets. The chemical behavior of the first of these ( $\text{SO}_2$ ) is normally that of a reducing agent in the narrow sense. The next two reagents ( $\text{S}^{--}$  and  $\text{CN}^-$ ), although they also frequently exhibit reductive reactivity, are characterized mainly by their tendency to coördinate with cations. The remaining three examples (Nos. 7 to 9) usually do not act as reducing agents in the restricted sense, but show their nucleophilic activity by combination with protons or other atomic nuclei.

The identical behavior of reagents such as the last five on the above list with respect to their union with a proton has previously been pointed out by J. N. Brönsted (5) and T. M. Lowry (26), who have founded a definition

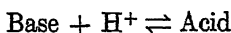
TABLE 8  
*Nucleophilic reagents*

NO.	REAGENT	NUMBER OF ELECTRONS NORMALLY INVOLVED IN REACTIONS	
		Completely donated	Share donated
1	$\text{Sn}^{++}$	2	—
2	$\text{Na}^+$ (metal)	1	—
3	$[\text{Fe}(\text{CN})]^{----}$	1	—
4	$\text{SO}_2$	—	2
5	$\text{CN}^-$	—	2
6	$\text{S}^{--}$	—	2, 4, 6, 8
7	$[\text{Co}(\text{H}_2\text{O})_5\text{OH}]^{++}$	—	2
8	$\text{NH}_3$	—	2
9	$\text{OH}^-$	—	2, 4

of the terms acid and base on the analogy between such equilibria as the following:



The defining equation is:



and a "base" and "acid" thus related are said to be "conjugate" with respect to each other. Evidently the broader classification includes the narrower; all "basic" reagents are nucleophilic, and "basicity" (affinity for a hydrogen nucleus) may be regarded as a special case of affinity for

atomic nuclei in general. Brönsted and Lowry pointed out that the differences in the states of electrification of reagents are trivial in comparison with the analogies of chemical behavior on which they based their definition, and the same statement obviously holds for the more general classification. Evidently nucleophilic activity is profoundly constitutive; it depends primarily on the presence of available electrons, and there is no simple and general connection with the gross electrification of the reagent.

### C. Electrophilic reagents

Reagents which acquire electrons, or a share in electrons, previously belonging to a foreign molecule or ion, may be termed *electrophilic* (15); table 9 contains some examples.

TABLE 9  
*Electrophilic reagents*

NO.	REAGENT	NUMBER OF ELECTRONS NORMALLY INVOLVED IN REACTION		FRAGMENTS OF REAGENT	
		Completely acquired	Share acquired	Liberated	Retained in product
1	$\text{S}_2\text{O}_8^{--}$	2	—	$\text{SO}_4^{--} + \text{SO}_4^{--}$	—
2	$[\text{Fe}(\text{CN})_6]^{---}$	1	—	} No splitting	
3	$\text{Fe}^{+++}$	1, 3	—		
4	$\text{O}_3$	—	2	$\text{O}_2$	} $\text{O}^{\bullet}$
5	$\text{SO}_5^{--}$	—	2	$\text{SO}_4^{--}$	
6	$\text{Cl}_2$	—	2	$\text{Cl}^-$	$\text{Cl}^+$
7	$\text{H}_3\text{O}^+$	—	2, 4	$\text{H}_2\text{O}$	} $\text{H}^+$
8	$\text{HBr}$	—	2	$\text{Br}^-$	
9	$\text{BF}_3$	—	2	} No splitting	
10	$\text{Be}^{++}$	—	8		

Corresponding to the subdivision of nucleophilic reagents already suggested, electrophilic reagents may be classified according to whether they act by appropriating electrons completely, or by acquiring only a share in electrons of foreign origin. The former category is illustrated by the first three examples in table 9; these reagents normally act by completely appropriating the number of electrons indicated, and they are all typical oxidizing agents in the sense in which the term is usually understood. The remaining seven reagents belong to the class which normally acts by acquiring a share in one or more electron duplets of external origin. The first two reagents of this series (Nos. 4 and 5) invariably behave as oxidizing agents in the restricted sense; the third (No. 6) does so frequently, but not invariably; the last four examples (Nos. 7 to 10) can be regarded as oxidizing agents only in the generalized sense, and indeed the whole list (Nos.

4 to 10) is so arranged that the connection with the restricted interpretation of oxidation becomes progressively less direct in the later examples.

In addition to this classification of electrophilic reagents, there is another natural method of subdivision, which is based on whether the valency electron shells are all complete or whether some are incomplete; for if all are complete, then new electrons can be accommodated only by the dissolution of a covalency, and the reagent must necessarily divide into two fragments; whereas if, on the contrary, certain of the electron shells are incomplete, these can accommodate the new electrons and the reagent will therefore cohere. (It is a corollary that electrophilic reagents with complete valency shells cannot be monatomic.) Referring, in the first place, to those reagents which act by acquiring electrons completely (see table 9, Nos. 1 to 3) the persulfate ion (No. 1) may be taken to illustrate the class in which all valency electron shells are complete; the acquisition of two additional electrons splits this reagent into two sulfate ions, which are both necessarily set free. The ferricyanide ion (No. 2) illustrates a polyatomic, and the ferric ion (No. 3) a monatomic, electrophilic reagent of the class characterized by the possession of an incomplete valency shell; such reagents acquire electrons without splitting. Turning next to the reagents (Nos. 4 to 10) which act by receiving only a share in new electrons, it will be evident that, as before, those reagents which already have complete valency shells must undergo fission; but there will be this difference in the present case, namely, that one of the fragments produced will be set free, the other having become bound in the product of reaction by the newly accepted electrons. Thus the permonosulfate ion leaves a neutral oxygen atom in combination and liberates a sulfate ion; again, the hydrogen bromide molecule leaves a proton in the reaction product and liberates a bromide ion; the table (Nos. 4 to 8) contains further illustrations. Reagents which acquire a share in new electrons and have incomplete valency shells do not undergo fission; on the contrary the whole reagent necessarily becomes included in the constitution of the product of the reaction. The table contains a polyatomic and monatomic example (Nos. 9 and 10) of reagents of this class, and their behavior may be illustrated by the formation of the borofluoride ion  $[\text{BF}_4]^-$ , and of the coordinated beryllium ion,  $[\text{Be}(\text{H}_2\text{O})_4]^{++}$ , respectively.

Reference may again be made to Brönsted's and Lowry's definition of acids and bases in order to indicate once more how the wider classification includes the more restricted; "acids" are those electrophilic reagents (see table 9, Nos. 7 and 8) which leave behind a proton in the product of reaction.

An inspection of table 9 as a whole shows that there is no simple connection between the state of electrification of a reagent and its electrophilic



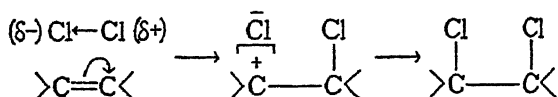
character; indeed it will be obvious with reference to reagents in general that their condition of electrification is a trivial matter in comparison with the analogies of behavior on the basis of which they are classified as nucleophilic or electrophilic as the case may be.<sup>5</sup>

It should be added that a certain number of reagents may be either nucleophilic or electrophilic according to circumstances; thus water is recognized both as a base and as an acid, since it can either gain a proton or lose one; and the nitrite ion exemplifies the type of reagent which can exhibit both the functions of reduction and oxidation.

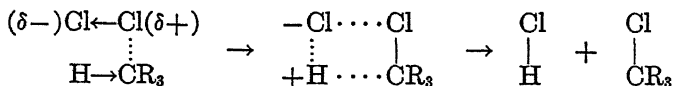
#### *D. Classification of some particular reagents*

It is proposed in this section to consider the electropolar character of certain commonly employed reagents in greater detail than was possible during the discussion of the principles underlying the classification of reagents in general.

The halogens and hypohalous acids are uniformly electrophilic, not only in their simple oxidizing and hydrolytic reactions, but also when they behave as additive and substituting agents in organic chemistry. When the chlorine molecule is polarized by the local electric field of a solvent molecule, or of a surface on which adsorption has occurred, it is the electron-depleted chlorine atom which may be expected to be the more reactive:  $(+\delta)\text{Cl} \rightarrow \text{Cl}(\delta-)$ . In an addition reaction under such conditions it may be supposed that the  $\text{Cl}^+$  is first bound with the rejection of  $\text{Cl}^-$ , which becomes accepted by the organic residue only as a consequence of the initial transaction (6):



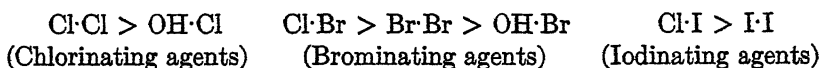
The case of substitution is rather simpler because  $\text{Cl}^+$  is all that is retained in the product of reaction, whilst  $\text{Cl}^-$  is liberated, usually in combination with a proton:



Once again the electron affinity of the polarized chlorine molecule may be held responsible for the initiation of reaction, although the proton affinity of the eliminated chloride ion is probably a facilitating influence (13, 19).

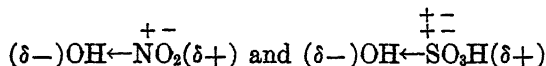
<sup>5</sup> Similarity to cations or anions in general being an unsound basis for the classification discussed, the categorization of reagents as "cationoid" or "anionoid" cannot be regarded as acceptable.

From the assumed electrophilic character of halogenating agents it follows that the reactivity of a compound  $X\cdot Cl$  in chlorination should increase with the electron affinity of  $X$ , and similarly for bromination and iodination; this agrees with the fact that bromine monochloride is a powerful brominating agent, but does not chlorinate, that similarly iodine monochloride iodates, and that generally the efficacy of halogenating agents is in the order which would be anticipated as a consequence of the electron affinity sequence  $Cl > Br > I > OH$ .

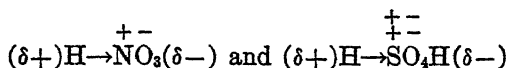


On the other hand, hypiodous acid is a more powerful iodinating agent than iodine, and this may mean that the proton affinity of the eliminated portion of the reagent cannot be neglected in the general assessment of relative reactivities.

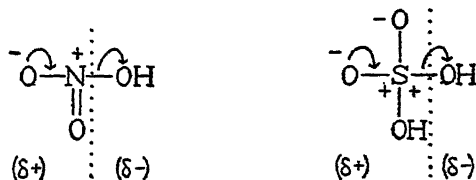
The behavior of nitric and sulfuric acids as substituting agents (13, 16) requires the assumption that they are electrophilic in consequence of polarizations



At first sight it seems remarkable that the polarizations do not correspond to the known modes of ionization. Obviously polarizations



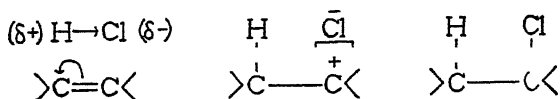
must occur; but the important point in the present connection is that they do not confer on the molecules the properties of a substituting agent and are therefore chemically unfruitful. The polarizations necessary for nitration and sulfonation, although doubtless assisted by the tautomeric electron displacements which both systems permit,



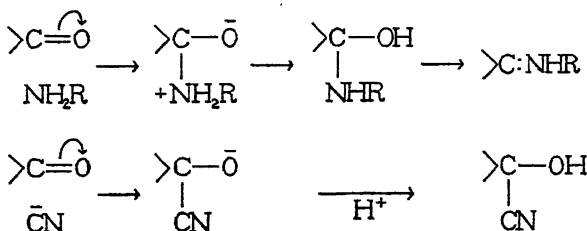
may be rare in time, and small in amplitude, except during the activation which accompanies reaction; but, when they do occur, they produce a positively charged atom with a depleted octet, and it is to the great electron

avidity of this part of the polarized structure that the properties of nitric and sulfuric acids as substituting agents must be ascribed.

The hydrogen halides are also classified as electrophilic, the proton being the reactive center, e.g.  $(\delta+)H \rightarrow Cl(\delta-)$ . Their addition to olefins, like the additions of the halogens themselves, are considered to be initiated by the binding of the positive atom (18):

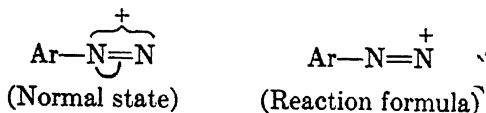


Ammonia, hydroxylamine, hydrazine and its derivatives, the cyanide ion, the bisulfite ion, and the anions of such compounds as ethyl malonate and ethyl acetoacetate, are all nucleophilic. They are the typical reagents for the carbonyl group, and the initiation of attack may in all cases be represented as the binding of an unshared electron pair by the nucleus of the carbonyl carbon atom (4, 7):

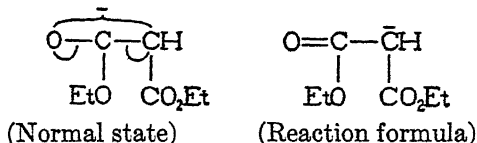


The condensations of such reagents as ethyl malonate with carbonyl compounds are known to be facilitated by catalysts (e.g., alcoholic sodium ethoxide) which would liberate the anion of the reagent, and in the case of the formation of cyanohydrins A. Lapworth has shown (23) that the rate-determining process is a reaction, not of hydrogen cyanide molecules, but of cyanide ions.

Reference may be made to the reactivities of mesomeric ions, since although these exist in forms which gain stability through degeneracy, they may be considered to react through their unperturbed structures. Thus, the electrophilic behavior of diazonium salts in coupling reactions and the like is a property of the cation, and is to be attributed to the circumstance that mesomerism creates a reactive center characterized by an incomplete valency shell:



Similarly, the reactivity in additions and condensations of such substances as nitromethane and the malonic esters is really to be ascribed in each case to the anion, which, on account of the circumstance that mesomerism produces carbon with unshared electrons, is highly nucleophilic:



Finally we may recall that electrophilic reagents stimulate electromeric polarizability effects of the  $+E$  type only, and nucleophilic reagents those of the  $-E$  type only, provided, in each case, that the system attacked admits of the appropriate form of polarizability.

#### REFERENCES

- (1) BAILEY, C. R., AND CASSIE, A. B. D.: *Proc. Roy. Soc. London* **140A**, 605 (1933).
- (2) BENNETT, G. M., AND GLASSSTONE, S.: *Proc. Roy. Soc. London* **145A**, 71 (1934).
- (3) BENNETT, G. M., AND MOSSES, A. N.: *J. Chem. Soc.* **1930**, 2364.
- (4) BLOOM, J., AND INGOLD, C. K.: *J. Chem. Soc.* **1931**, 2765.
- (5) BRÖNSTED, J. N.: *Rec. trav. chim.* **42**, 718 (1923).
- (6) BURTON, H., AND INGOLD, C. K.: *J. Chem. Soc.* **1928**, 910.
- (7) COOPER, K. E., INGOLD, C. K., AND INGOLD, E. H.: *J. Chem. Soc.* **1926**, 1868.
- (8) FLÜRSCHHEIM, B.: *J. prakt. Chem.* **66**, 321 (1902); *J. Chem. Soc.* **95**, 722 (1909); **97**, 91 (1910); *J. Soc. Chem. Ind.* **44**, 246 (1925).
- (9) FRY, H. S.: *The Electronic Conception of Valence*. Longmans, Green and Co., London and New York (1921).
- (10) HÖJENDAL, K.: *Studies of Dipole Moments*. Bianco Lunos Bogtrykkeri, Copenhagen (1928).
- (11) INGOLD, C. K.: *Ann. Repts. Chem. Soc.* **23**, 129 (1926).
- (12) INGOLD, C. K.: *Ann. Repts. Chem. Soc.* **23**, 144 (1926).
- (13) INGOLD, C. K.: *Ann. Repts. Chem. Soc.* **24**, 153 (1927).
- (14) INGOLD, C. K.: *Rec. trav. chim.* **48**, 797 (1929).
- (15) INGOLD, C. K.: *J. Chem. Soc.* **1933**, 1120.
- (16) INGOLD, C. K., AND INGOLD, E. H.: *J. Chem. Soc.* **1926**, 1310.
- (17) INGOLD, C. K., AND SHAW, F. R.: *J. Chem. Soc.* **1927**, 2918.
- (18) INGOLD, C. K., AND SMITH, H. G.: *J. Chem. Soc.* **1931**, 2752.
- (19) INGOLD, C. K., SMITH, E. W., AND VASS, C. C. N.: *J. Chem. Soc.* **1927**, 1245.
- (20) INGOLD, C. K., AND VASS, C. C. N.: *J. Chem. Soc.* **1928**, 417.
- (21) LANGMUIR, I.: *J. Am. Chem. Soc.* **42**, 274 (1920).
- (22) LAPWORTH, A.: *J. Chem. Soc.* **73**, 445 (1898); **79**, 1265 (1901); *Proc. Manchester Phil. Soc.* **64**, No. 3 (1920); *J. Chem. Soc.* **121**, 416 (1922).
- (23) LAPWORTH, A.: *J. Chem. Soc.* **83**, 995 (1903); **85**, 1206, 1214 (1904); **89**, 945 (1906).
- (24) LEWIS, G. N.: *J. Am. Chem. Soc.* **35**, 1448 (1913); *Proc. Nat. Acad. Sci.* **2**, 586 (1916). *Valence and the Structure of Atoms and Molecules*. The Chemical Catalog Co., Inc., New York (1923).
- (25) LEWIS, G. N.: *Valence and the Structure of Atoms and Molecules*, especially p. 139. The Chemical Catalog Co., Inc., New York (1923).

- (26) LOWRY, T. M.: J. Soc. Chem. Ind. **42**, 43 (1923).
- (27) LOWRY, T. M.: J. Chem. Soc. **123**, 822, 1886 (1923); *Nature* **114**, 376 (1925).
- (28) LUCAS, H. J., ET AL.: J. Am. Chem. Soc. **46**, 2475 (1924); **47**, 1459, 1462 (1925).
- (29) LÜTGERT, H.: Z. physik. Chem. **14B**, 31 (1931).
- (30) MICHAEL, A.: J. prakt. Chem. **46**, 204 et seq. (1892); **60**, 286 et seq. (1899).
- (31) PARTS, A.: Z. physik. Chem. **4B**, 227 (1929).
- (32) PAULING, L.: J. Chem. Physics **1**, 280 (1933).
- (33) PAULING, J., AND SHERMAN, J.: J. Chem. Physics **1**, 606 (1933).
- (34) PAULING, L., AND SHERMAN, J.: J. Chem. Physics **1**, 679 (1933).
- (35) PAULING, L., AND WHEELAND, G. W.: J. Chem. Physics **1**, 362 (1933).
- (36) ROBINSON, R., ET AL.: J. Chem. Soc. **1926**, 401.
- (37) SCHAEFER, C.: Trans. Faraday Soc. **24**, 841 (1929).
- (38) SIDGWICK, N. V.: Chem. Rev. **9**, 77 (1931).
- (39) SIDGWICK, N. V., SUTTON, L. E., AND THOMAS, W.: J. Chem. Soc. **1933**, 406.
- (40) STIEGLITZ, J.: J. Am. Chem. Soc. **44**, 1293 (1922).
- (41) SUTTON, L. E.: Proc. Roy. Soc. London **133A**, 668 (1931).
- (42) THIELE, J.: Ann. **306**, 87 (1899).
- (43) THOMSON, J. J.: Phil. Mag. **46**, 513 (1923).
- (44) WILLIAMS, J. W.: Physik. Z. **29**, 174, 683 (1928).

# THE HALOFORM REACTION

REYNOLD C. FUSON AND BENTON A. BULL

*Department of Chemistry, University of Illinois, Urbana, Illinois*

*Received September 28, 1934*

## CONTENTS

I. Introduction.....	275
II. The early history of the haloform reaction.....	276
III. The haloform reaction in qualitative organic analysis.....	277
IV. Structural determination by means of the haloform degradation.....	281
V. Quantitative methods based on the haloform reaction.....	286
VI. The use of the haloform reaction in synthesis.....	287
A. The haloforms.....	287
B. Saturated aliphatic acids.....	288
C. Unsaturated acids.....	289
D. Aromatic acids.....	290
VII. The mechanism of the haloform reaction.....	291
A. The halogenation phase of the haloform reaction.....	292
B. The cleavage phase of the haloform reaction.....	295
VIII. Reactions related to the haloform reaction.....	298
A. Modifications of the haloform reaction.....	299
B. Reactions analogous to the haloform reaction.....	299
IX. Summary.....	304
X. References.....	304

## I. INTRODUCTION

The haloform reaction comprises those processes whereby the haloforms are derived from organic compounds by the action of hypohalites. It represents one of the most interesting and useful types of organic reactions and, since its discovery more than a century ago, has engaged the attention of investigators in a wide variety of ways. The theoretical problem involved is unique; and, in spite of the large amount of study which has been expended upon its various aspects, the mechanism of the process is not yet fully understood. On the practical side, the haloform reaction has furnished indispensable methods of detection, estimation, synthesis, and degradation. The literature of the subject is extensive, but is made up chiefly of observations incidental to other studies. This review represents the first general survey which has been made of the work in this field. The fact that many of the examples of the haloform reaction have been carried out incidentally in connection with other work has made the literature

survey peculiarly difficult. No attempt has been made to present here a complete bibliography, but the most significant contributions to the subject have been cited.

## II. THE EARLY HISTORY OF THE HALOFORM REACTION

The history of the haloform reaction dates from 1822, when Serullas (167) made the accidental discovery that iodine and alkali convert alcohol into iodoform, which he called a "hydroiodide of carbon." Chloroform was not known until Guthrie (72) in 1831 isolated an "impure chloric ether" formed by the action of bleaching powder on alcohol, and Soubeiran (173) prepared "bichloric ether" by the same method. This compound was discovered almost simultaneously by these investigators and by Liebig (126), who prepared it by the action of alkali on chloral, under circumstances which led to the celebrated dispute regarding priority in this connection—a dispute which was carried on throughout the life-time of these investigators (128), and which is still the subject of much discussion (37, 42, 174, 30).

In 1834, Dumas (44) prepared bromoform by the action of hypobromite on alcohol. He determined the composition of the haloforms and gave them their present names, which refer to the fact that they yield formic acid on hydrolysis.

In extending the reaction Dumas and Peligot (45) reported the formation of chloroform by the action of hypochlorite on methyl alcohol; Lefort (121) also claimed that this alcohol gave the iodoform test. Bonnet (19) asserted that acetic acid gave chloroform when treated with bleaching powder. The error with methyl alcohol occurred repeatedly in the literature until acetone-free methanol became available in comparatively recent times. As a matter of fact, Lieben (125) showed in 1870 that pure methyl alcohol, obtained by hydrolysis of dimethyl oxalate, did not give iodoform,—a result which was confirmed by Belohoubek (11) in 1873.

The reports contained in the early literature did not afford any useful generalization. Boucharadat (21), using hypoiodite, reported positive tests with ethyl acetate and ethyl ether, and a negative test with methyl ethyl ketone, while Schlagdenhauffen (158) reported the formation of chloroform from the ethyl esters of formic, acetic, tartaric, benzoic, and nitrous acids, as well as from methyl tartrate. Likewise, Millon (134) in 1845 reported the formation of iodoform when sucrose, glucose, gums, dextrin, or albuminoids were treated with iodine and potassium bicarbonate.

It was not until 1870 that any systematic study of the reaction was made. In this year Lieben (124) showed that the "iodoform test" was given by a wide variety of compounds, including acetaldehyde, ethyl alcohol, acetone, ethyl acetate, *secondary*-butyl alcohol, acetophenone, 2-hexanone, lactic acid, and capryl alcohol. He also reported positive tests with amylene,

butyraldehyde, *n*-butyl alcohol, quinic acid, dulcin, sugars, meconic acid, methyl butyrate, propionaldehyde, and *n*-propyl alcohol; in these cases we now believe the production of iodoform must have been due to the presence of impurities such as alcohol and acetone. He proved, however, that pure ether and pure acetic acid do not respond to the test. He also reported negative tests for benzene, ethyl chloride, ethyl bromide, ethylene dichloride, amyl alcohol, benzaldehyde, anisic acid, benzoic acid, malic acid, formic acid, succinic acid, butyric acid, chloral hydrate, carbon tetrachloride, citric acid, glycerol, glycol, urea, oxalic acid, phenol, and picric acid.

From this study Lieben announced a general rule regarding the significance of this reaction with respect to molecular structure. The rule which is used at the present time is essentially the same as that of Lieben, and may be stated as follows: *A positive iodoform test is given by compounds containing the aceto ( $\text{CH}_3\text{CO}$ —) group joined to either carbon or hydrogen, and by compounds which are oxidized under the conditions of the test to derivatives containing this structural unit.*

That this rule is subject to considerable revision is indicated by the results of subsequent experiments, not only with hypiodite but with the other hypohalites as well. Thus, with hypiodite certain acetylides (118) yield iodoform; and with hypobromite, dihydroresorcinol (196), citric acid (26), certain olefins and tertiary alcohols (179), and many sugars and alkaloids (32) are reported to yield bromoform or carbon tetrabromide.<sup>1</sup> The formulation of a more general rule has been accomplished in connection with the use of Lieben's test in qualitative organic analysis.

### III. THE HALOFORM REACTION IN QUALITATIVE ORGANIC ANALYSIS

Since 1870, when Lieben laid down the general rule regarding the structural implications of the haloform reaction, the "iodoform test" has become one of our most useful tools in the identification of alcohols and ketones which are soluble in water. The original procedure involved the treatment of the compound to be tested with iodine in an excess of dilute aqueous alkali. In some cases it was advisable to heat the mixture, and sometimes long standing was advantageous.

That this procedure is widely applicable is evident from an examination of the following list of compounds, all of which have been shown to give a positive iodoform test:<sup>2</sup> ethyl alcohol (124), acetaldehyde (124,

<sup>1</sup> It is probable that the carbon tetrabromide in these cases is formed by the action of the hypobromite on bromoform (32).

<sup>2</sup> As has been indicated, exceptions to the Lieben rule are not unknown. Thus, acetoacetic acid (35) does not give the iodoform test. It should be mentioned in this connection that Hurd and Thomas found that acetoacetic ester with bleaching



112),<sup>3</sup> acetone (124), isopropyl alcohol (112, 190), 2-butanone (112, 35), *sec*-*ondary*-butyl alcohol (124), 2-pentanone (35), 2-hexanone (124, 35), 2-octanol (124), 2-octanone (112, 35), acetophenone (124, 112, 35), levulinic acid (190), pyruvic acid (35), lactic acid (124, 190), mesityl oxide (34), 6-ethyl-3,4-dimethoxyacetophenone (168), 6-ethyl-3-ethoxy-4-methoxyacetophenone (109), 1-diethylaminobutanone-3 (172), and certain methyl sugars (14).

For water-insoluble compounds, however, the Lieben test is uncertain and likely to be misleading. A more useful form of the test has been developed by Fuson and Tullock (62), who found that dioxan could be used as a solvent. This modified test is carried out by dissolving the compound to be tested in 5 cc. of dioxan, adding dilute aqueous alkali, and a slight excess of iodine-potassium iodide solution, and then warming the mixture for two minutes at 60°C. After the excess iodine is discharged by means of alkali, the iodoform is precipitated by the addition of water.

The new procedure has been tried with a large number of compounds with results as indicated below.

#### *Alcohols*

*Positive:* isopropyl alcohol, methyl-*n*-amylcarbinol, 2-octanol, methylisopropylcarbinol, 2,3-butanediol, methylbenzylcarbinol.

*Negative:* methyl alcohol, allyl alcohol, trimethylene glycol, mannitol, isobutyl alcohol.

#### *Aliphatic ketones*

*Positive:* acetone, methyl ethyl ketone, methyl propyl ketone, 2-hexanone, methyl isobutyl ketone, 2-heptanone, 2-octanone, methyl isohexyl ketone, 4-methyl-2-heptanone, methyl cyclohexyl ketone, methyl  $\gamma$ -phenoxypropyl ketone, benzylacetone, benzohydrilacetone.

#### *Mixed ketones*

*Positive:* acetophenone, methyl *p*-tolyl ketone, *p*-chloroacetophenone, *p*-bromoacetophenone, methyl *p*-anisyl ketone, 2,4-dimethoxyacetophenone, 2-methyl-4-methoxyacetophenone, 5-methyl-2-methoxyaceto-

---

powder gave dichloroacetic acid in a yield of 60 per cent (94). Pinacolone (39) and many similar compounds involving steric hindrance fail to respond to the Lieben test. On the other hand, many compounds which do not possess either of the groupings  $\text{CH}_3\text{COC}-$  and  $\text{CH}_3\text{CH}(\text{OH})\text{C}-$  give a positive Lieben test. Examples of this type are acetoxime, 2-methyl-2-butene, and pulegone (147).

<sup>3</sup> It is interesting to note that the corresponding acetal,  $\text{CH}_3\text{CH}(\text{OC}_2\text{H}_5)_2$ , does not give a positive iodoform test. However, if the acetal is shaken with a drop of hydrochloric acid it will then give the test (71).

phenone, acetocymene, 2,4,5-trimethylacetophenone, *o*-hydroxyacetophenone, *m*-hydroxyacetophenone, *p*-hydroxyacetophenone, 3-methoxy-4-hydroxyacetophenone, *o*-nitroacetophenone, *m*-nitroacetophenone, *p*-nitroacetophenone, *o*-aminoacetophenone, *m*-aminoacetophenone, *p*-aminoacetophenone, 2-aceto-1-naphthoxyacetic acid, 2-aceto-4-bromo-1-naphthoxyacetic acid.

*Negative:*  $\alpha$ -chloroacetophenone, propiophenone, acetomesitylene, 3,5-dinitroacetomesitylene, 2,4,6-tribromoacetophenone, 3-amino-2,4,6-tribromoacetophenone, 1-aceto-2-naphthoxyacetic acid, 2-methoxy-1-acetonaphthone.

#### *Unsaturated ketones*

*Positive:* mesityl oxide, benzalacetone, 2-methyl-1-phenyl-1-buten-3-one, furfuralacetone.

#### *Diketones*

*Positive:* acetylacetone, acetonylacetone, benzoylacetone, *p*-bromobenzoylacetone, dibenzoylmethane, 1,3-diketohydrindene, 2,6-dimethyl-4-acetylacetophenone.

*Negative:*  $\omega$ -acetylacetomesitylene,  $\omega$ -benzoylacetomesitylene, di( $\beta$ -isoduryloyl)methane.

#### *Acids and acid derivatives*

*Positive:* ethyl lactate,  $\alpha$ -aminoisobutyric acid, levulinic acid, ethyl levulinate,  $\gamma$ -acetylbutyric acid, diethyl acetylsuccinate, diethyl  $\alpha, \alpha'$ -diacetylsuccinate.

*Negative:*\* alanine,\* secondary-butyl acetate,\* secondary-amyl acetate,\* diethyl phthalate,\* diethyl adipate.\*

#### *Miscellaneous*

*Positive:* acetoxime, diacetyl monoxime,  $\alpha$ -phenylethylamine.

*Negative:* 2-pentene, 1,1-diphenyl-1-propene, 1-chloro-2,3-dihydroxypropane, propionitrile, isoeugenol, phenylacetylene, rhamnose, acetophenone oxime,\* anethole, resorcinol, phloroglucinol.

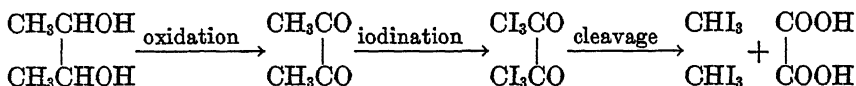
The foregoing results of Fuson and Tullock were summarized in the following rule: *The test is positive for compounds which contain the grouping  $\text{CH}_3\text{CO}$ —,  $\text{CH}_2\text{ICO}$ —, or  $\text{CHI}_2\text{CO}$ —<sup>5</sup> when joined to a hydrogen atom or to a carbon atom which does not carry highly activated hydrogen atoms or groups*

<sup>4</sup> Compounds marked with an asterisk give iodoform, but require a longer period of heating than that specified in the procedure given.

<sup>5</sup> Compounds containing the  $\text{Cl}_3\text{CO}$ — group would undoubtedly give a positive test also, but as yet no such compound is known in a pure state.

which provide an excessive amount of steric hindrance. The test will, of course, be positive also for any compound which reacts with the reagent to give a derivative containing one of the requisite groupings. Conversely, compounds which contain one of the requisite groupings will give a negative test in case this grouping is destroyed by the hydrolytic action of the reagent before iodination is complete.

A theoretical basis for the foregoing generalization is to be found in the three fundamentally different types of effects produced by the reagent. It is capable of oxidizing alcohols and amines, of replacing active hydrogen atoms by iodine atoms, and of cleaving certain types of carbon chains. For example, these three types of processes undoubtedly take place in the order mentioned when 2,3-butanediol is converted to iodoform:



It is important to note, however, that the iodination will occur, not necessarily on an active methyl group, but rather at the point in the molecule where the most active hydrogen atoms are found. Thus,  $\beta$ -diketones having the grouping  $\text{—COCH}_2\text{CO—}$  are iodinated first at the methylene group. Apparently in these cases chain cleavage then occurs. The soundness of this explanation is attested by the fact that dibenzoylmethane and 1,3-diketohydrindene give a positive test; here the iodoform obviously derives from the methylene group, and diiodomethyl ketones—compounds containing one of the requisite groupings—must be postulated as intermediates. Similarly, although acetylacetone, benzoylacetone, and *p*-bromobenzoylacetone contain the acetyl group, the iodoform comes rather from the methylene group. This conclusion is, in turn, supported by the fact that  $\omega$ -acetylacetomesitylene ( $\text{CH}_3\text{COCH}_2\text{COC}_6\text{H}_4$ ) gives a negative test; this is to be expected, since diiodoacetomesitylene, on account of the hindrance involved, would not yield iodoform (147, 60).

Numerous compounds which do not contain one of the requisite groupings give a positive test presumably because, under the influence of the reagent, they give rise to products which do contain such groupings. Many alcohols are oxidized to the corresponding carbonyl compounds and so lead to the formation of iodoform. Amines of similar structures apparently undergo analogous transformations. In the case of certain esters and oximes the production of iodoform must be traced to hydrolysis products.

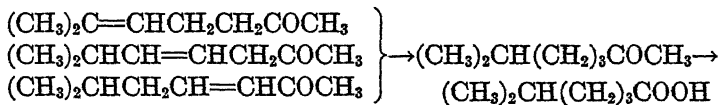
The iodoform reaction is greatly retarded by steric hindrance. The test is negative for all compounds which contain one of the requisite groupings joined to an aryl radical carrying two ortho substituents. The question as to what is actually formed in the case of hindered methyl ketones has

been solved by Poggi (147), who has shown very recently that when the procedure of Lieben is used pinacolone does not give iodoform<sup>6</sup> but is converted to diiodopinacolone,  $(\text{CH}_3)_2\text{CCOCHI}_2$ . Similar results have been obtained by Fuson, Johnson, and Bull, who isolated diiodoacetomesitylene ( $\text{C}_9\text{H}_{11}\text{COCHI}_2$ ) and 3-diiodoacetyl-2,4,6-trimethylbenzoic acid ( $\text{HOOC-C}_9\text{H}_9\text{COCHI}_2$ ) by the action of hypiodite on acetomesitylene and 3-acetyl-2,4,6-trimethylbenzoic acid, respectively (60).

#### IV. STRUCTURAL DETERMINATION BY MEANS OF THE HALOFORM DEGRADATION

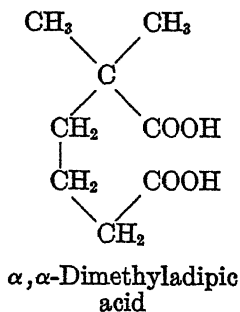
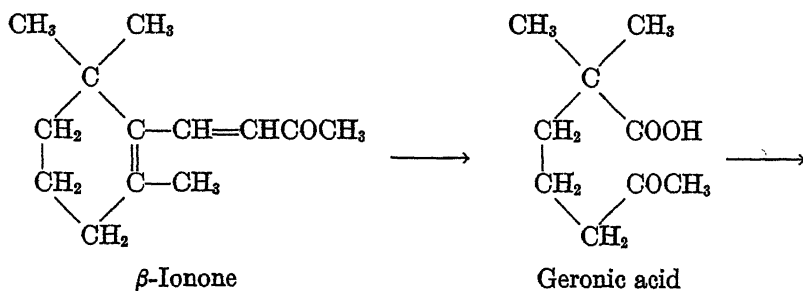
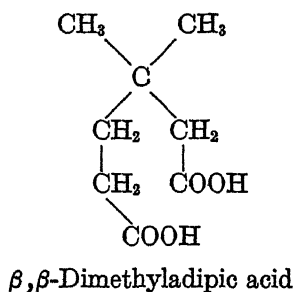
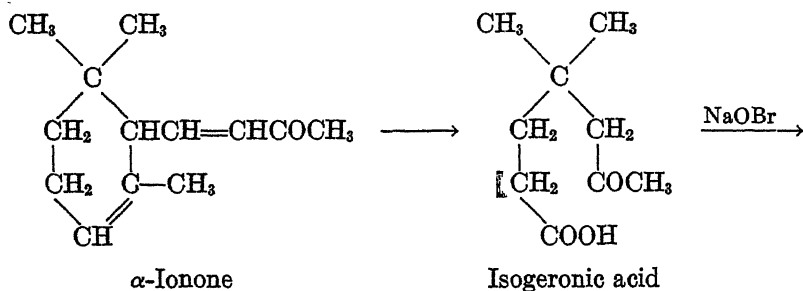
The great utility of the haloform reaction in the proof of structure depends primarily on the fact that it effects simultaneously the detection and oxidative degradation of a methyl ketone group. Moreover, it is highly specific for this grouping, and as a consequence can often be used safely with compounds which would be profoundly altered by ordinary oxidizing agents. It is impracticable to tabulate all of the compounds whose structures have been elucidated with the aid of the haloform reaction; several examples of its application to the field of terpene chemistry, wherein it has been of especial value, have been selected to illustrate its wide use as a degradative procedure.

The methyl heptenones, together with isoprene, are commonly regarded as key substances in the chemistry of the terpenes, not only because they are frequently isolated in degradative studies, but also because they are often the basis of synthetic work. The relation between the three most common methyl heptenones was demonstrated by Wallach, who found that they were all reduced to the same heptanone, and that the latter compound gave isoheptoic acid when treated with sodium hypobromite (207):

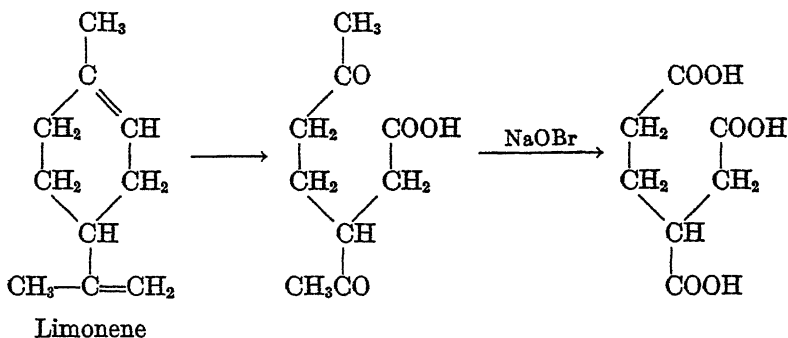


The haloform degradation has played a part in the structural studies of nearly every type of mono- and di-cyclic terpene. The ionones may be cited as typical examples. The positions of the double bonds in these compounds were established by the brilliant investigations of Tiemann, who oxidized the ionones to the corresponding ketonic acids, which were then converted by means of the haloform reaction to dimethyladipic acids (187, 186):

<sup>6</sup> The course of the reaction in the case of pinacolone is apparently dependent on the conditions, for with the procedure of Fuson and Tullock iodoform is obtained (62).

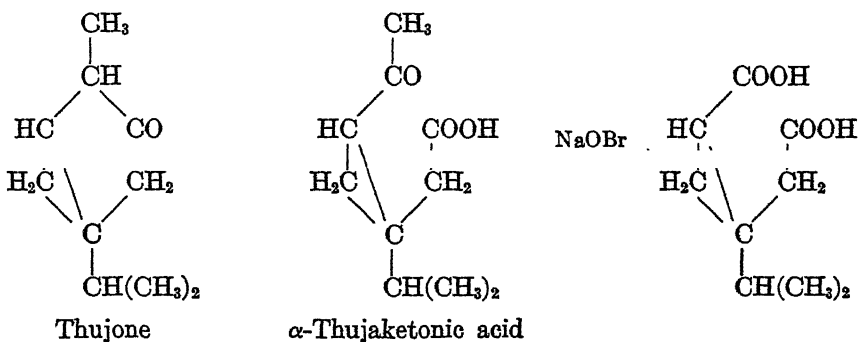


In a similar manner the structure of limonene (dipentene) was established by Tiemann and Semmler (188):

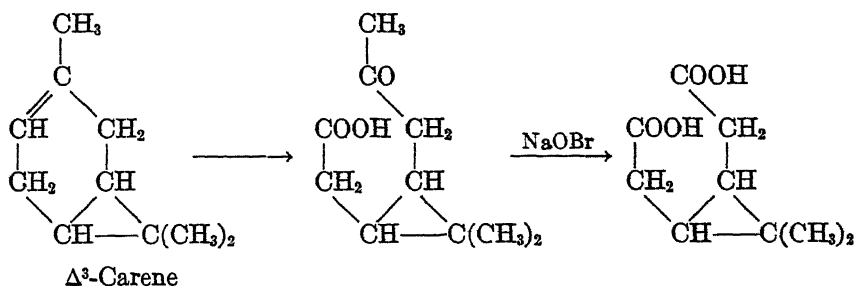


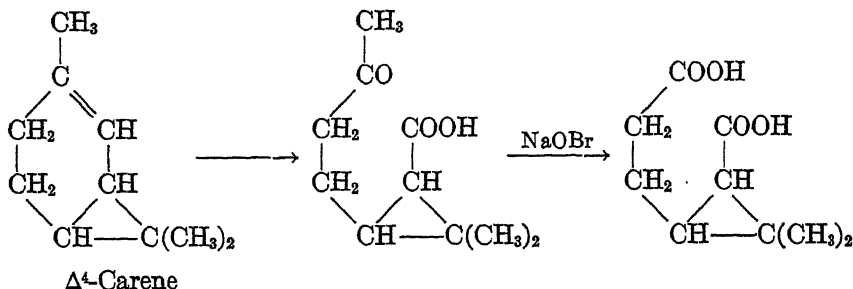
The  $\beta,\delta$ -diacetylvaleric acid obtained by oxidation of limonene is converted by sodium hypobromite into the corresponding tribasic acid.

A similar degradation was involved in the chain of evidence which enabled Semmler (162) to arrive at the correct structure of thujone. Oxidation of thujone gave  $\alpha$ -thujaketonic acid and this was converted to the cyclopropanedicarboxylic acid by means of hypobromite:

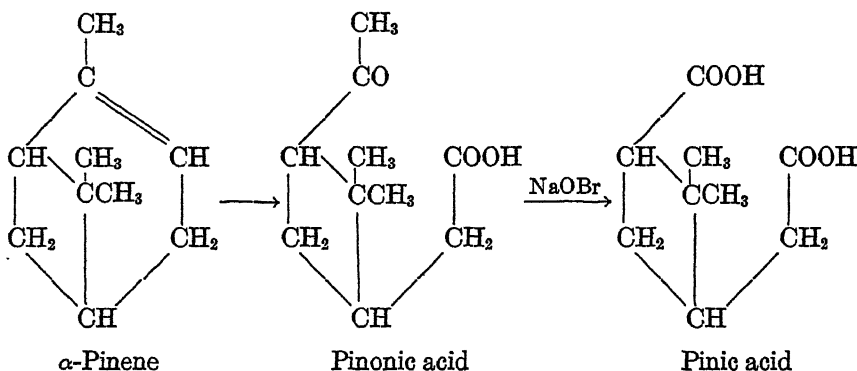


In the carane series, the positions of the double bonds in  $\Delta^3$ -carene (166) and  $\Delta^4$ -carene (170) were determined by oxidation, followed by treatment with hypobromite:

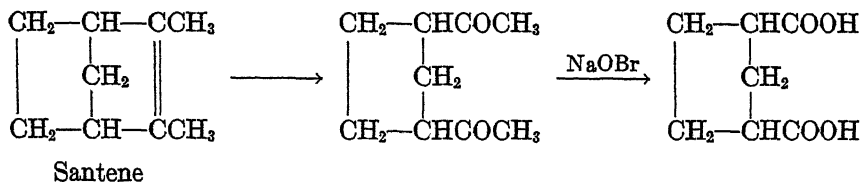




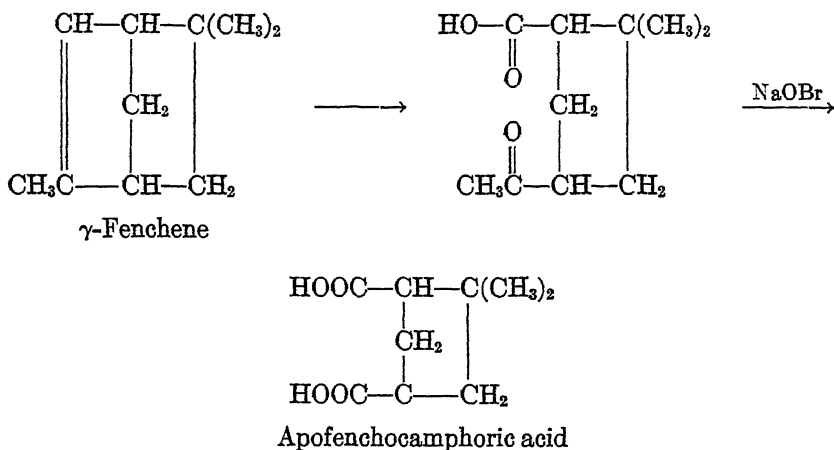
A similar procedure was used in the classical investigation by Baeyer of the structure of the pinanes. The pinonic acid resulting from the oxidation of  $\alpha$ -pinene with permanganate was degraded to pinic acid by means of hypobromite (5, 7).



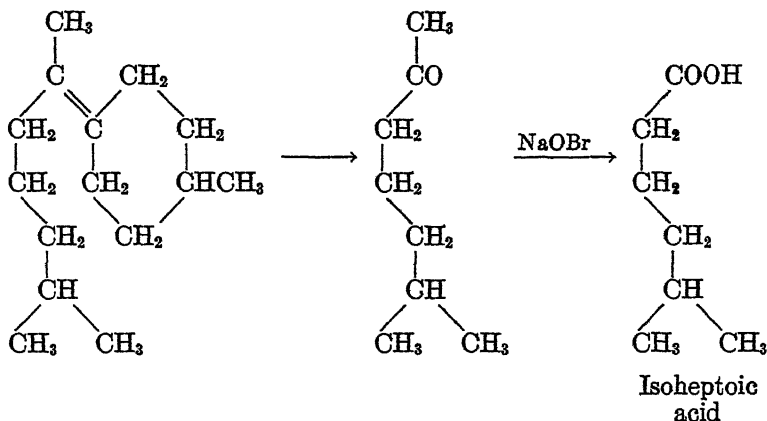
Semmler and Bartelt determined the structure of santene by ozonization followed by treatment with hypobromite (163):



$\gamma$ -Fenchene of the fenchane series, when subjected to ozonolysis, gave a keto aldehyde and a keto acid; the latter, when treated with hypobromite, was converted into apofenchocamphoric acid (107):

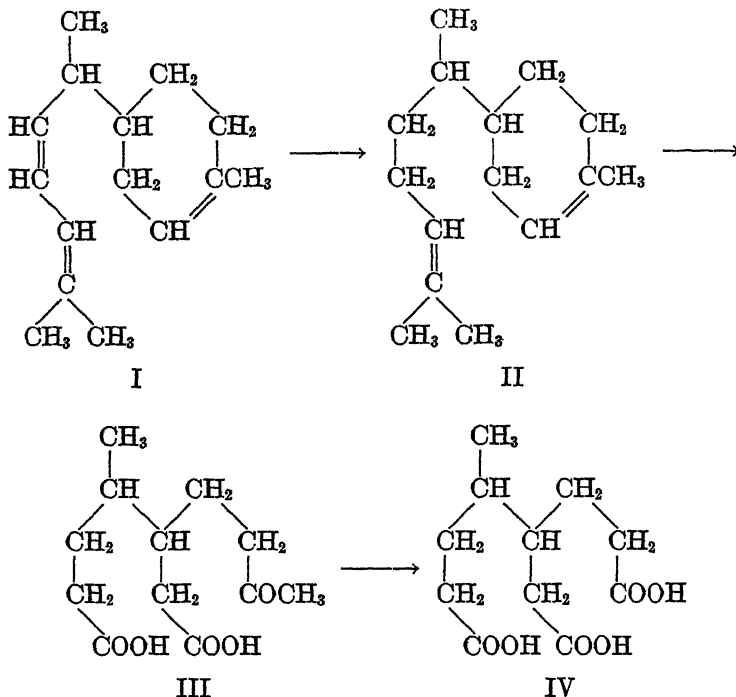


Investigators of the sesquiterpenes have repeatedly used the haloform reaction but, in most cases, it is too early to estimate the significance of this work, since the structural problems involved are not entirely solved. Mention may be made, however, of the work of Ruzicka and van Veen, who oxidized tetrahydrobisabolene and then treated the acyclic ketonic fragment with hypobromite, which gave isoheptoic acid (155):



Zingiberene (I) may be mentioned also, as the haloform degradation has been particularly useful in the elucidation of its structure (156). Partial reduction followed by oxidation with permanganate converts zingiberene into a keto acid (III) which, with hypobromite, gives the corresponding tribasic acid (IV):





#### V. QUANTITATIVE METHODS BASED ON THE HALOFORM REACTION

The fact that iodoform is a solid which is easy to isolate has led to many attempts to use the iodoform reaction in the quantitative estimation of compounds which undergo the reaction. As early as 1870 Lieben, by this means, was able to detect alcohol in water in dilutions as high as 1 part in 2000 (124). Krämer (112) subsequently developed a gravimetric procedure based on the iodoform reaction, in which the iodoform was extracted with ether and weighed after the evaporation of the volatile solvent. This original gravimetric method has been used by several investigators in the estimation of acetone in the presence of methanol (82, 2, 194, 104).

An excellent volumetric method was developed by Messinger (133, 195), which consisted in liberating the excess iodine used in the formation of iodoform, absorbing the free halogen in excess thiosulfate, and back-titrating with iodine, using starch as an indicator. Various modifications of this method have been suggested for the estimation of acetone in the presence of water (33, 66, 80, 148, 69), methanol (152, 148, 69, 10), ethanol (152, 175, 219, 148, 69, 105), urine (85), or glycerol (12). In addition, Elliott and Dalton used Messinger's iodometric method for estimating the amount of acetone in air (48). Hatcher and Mueller (79) state, however,

that with pyruvic acid, acetaldehyde, and acetone, sodium hypoiodite does not give a quantitative yield of iodoform under any conditions. The Mes-singer method has been studied also by Kebler (99), Vaubel and Scheuer (193), Keppeler (101), and Marriott (131).

van der Lee (119) found that the formation of iodoform from acetone takes place at a much lower pH than from alcohol. Thus, in testing acetone, sodium carbonate may be used, but with ethyl alcohol caustic alkali is necessary. Similarly, levulinic acid reacts more rapidly than does isopropyl alcohol or lactic acid. Upon this fact Kolthoff (105) has based a very sensitive test for acetone in the presence of alcohol by control of the pH used. An extremely sensitive form of this test was discovered by Kunz (115), who found that the presence of potassium persulfate imparted a red or orange color to the precipitated iodoform. Ammonium persulfate may also be used (161). With this reagent the formation of iodoform may be detected in concentrations as low as 1 part in 200,000 by the turbidity which appears.

The haloform reaction has also been employed in the quantitative estimation of lactic acid (96), levulinic acid (157), and pyruvic acid (79).

## VI. THE USE OF THE HALOFORM REACTION IN SYNTHESIS

### A. *The haloforms*

The synthesis of the haloforms themselves has almost always been accomplished by the action of hypohalites on ethyl alcohol or acetone. Chloroform, the most important commercially, is generally prepared by the action of bleaching powder on alcohol. Bromoform and iodoform, on the other hand, are best prepared by the electrolytic process (68). Iodoform is best prepared electrolytically from alcohol (181, 117) and bromoform from acetone, while the corresponding method of preparing chloroform is uneconomical (68).<sup>7</sup>

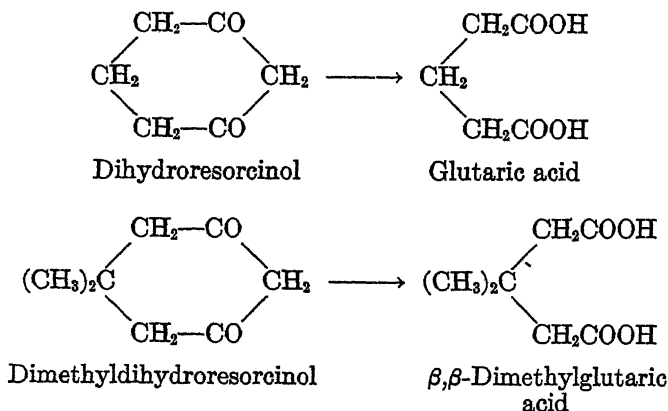
The haloforms, when prepared by the haloform reaction, are contaminated to a greater or less degree by the corresponding tetrahalomethanes. This is due to the action of the hypohalites on the haloforms by which the latter are converted into the corresponding tetrahalogen compounds:  $\text{CHX}_3 + \text{NaOX} = \text{CX}_4 + \text{NaOH}$ . This is a general reaction (83, 73, 8, 38); it often becomes the principal reaction when dilute solutions are used. Thus, Wallach (198) found that in the formation of bromoform from hypobromite and acetone in very dilute solutions, the reaction was slow and that carbon tetrabromide was the chief product. This influence of the pH on the course of reactions involving hypohalites is frequently noticed (105, 119, 79).

<sup>7</sup> For excellent reviews of the electrolytic methods, see Feyer (53) and Glass-tone (68).

*B. Saturated aliphatic acids*

The saturated aliphatic acids are occasionally prepared by means of the haloform degradation of alcohols and ketones. The formation of trimethylacetic acid from pinacolone by the action of potassium hypochlorite (17) or potassium hypobromite (151, 67) is a synthesis of this type. Similarly, *tertiary*-butylacetic acid is made by the action of sodium hypobromite on 4,4-dimethyl-2-pentanone (86). A very large number of similar preparations is known, but in most cases the acid is more readily obtainable than the corresponding secondary alcohol or methyl ketone, and, as a consequence, the method has limited value in the synthesis of the fatty acids. Among the monobasic acids which have been made in this way are  $\alpha$ , $\alpha$ -diphenylbutyric (132),  $\alpha$ -methyl- $\alpha$ -ethylbutyric (141),  $\beta$ -phenylisovaleric (84),  $\delta$ -phenylvaleric (20),  $\alpha$ , $\alpha$ -dimethylcaproic (120), isoeptioic (207), caprylic (183), and capric (183, 74). Methyl ketonic acids are frequently encountered—especially among the oxidation products of terpenes—and in many instances these have been further degraded to the corresponding dibasic acids by means of the haloform method. Examples of these acids are methylsuccinic (145), isopropylsuccinic (199), isobutylsuccinic (13),  $\alpha$ , $\alpha$ -dimethylsuccinic (185, 102, 210, 116), trimethylsuccinic (129),  $\alpha$ , $\alpha$ -dimethylglutaric (184),  $\alpha$ -isopropylglutaric (199, 209, 164),  $\beta$ -isopropylglutaric (22, 201),  $\alpha$ -methyl- $\beta$ -isopropylglutaric (199),  $\beta$ -methyladipic (206),  $\alpha$ -isopropyladipic (212),  $\beta$ -isopropyladipic (202),  $\alpha$ , $\alpha$ -dimethyladipic (187), pimelic (123, 200), and  $\beta$ , $\beta$ -dimethylpimelic (122, 211). Lactonic acids have been made from the lactones of certain  $\delta$ -acetyl- $\gamma$ -hydroxyvaleric acids (81).

Of especial interest is the formation of glutaric acid from dihydroresorcinol (196) and of  $\beta$ , $\beta$ -dimethylglutaric acid from dimethyldihydroresorcinol or its monobromo derivative (106, 196):

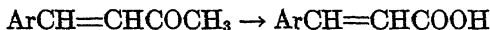


The haloform degradation appears to be generally applicable to diketones of this type.

As has been indicated in section IV, many alicyclic acids have also been made by the haloform degradation. Other examples are hexahydrobenzoic acid (204), the hexahydrotoluic acids (165, 213, 153), cyclopentanecarboxylic acid (138) and certain of its homologues (203, 139), norpinic acid (197, 103), cyclohexylacetic acid (205) and hexahydrotolylacetic acids (153),  $\beta$ -(2,2',6-trimethylcyclohexyl)propionic acid (154), and 3-carboxy-1,1-dimethylcyclopropane-2-propionic acids (170, 144).

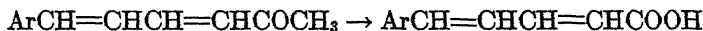
### C. Unsaturated acids

Perhaps the most elegant application of the haloform reaction is in connection with unsaturated methyl ketones with which the usual oxidizing agents cannot be used. Hypohalites convert the acetyl group into a carboxyl group without altering the rest of the molecule. An early patent (46) sets forth a method for preparing cinnamic acid and its nuclear substitution products by the action of hypohalites on the corresponding benzalacetones:



$\alpha$ -Methyl-*p*-isopropylcinnamic acid (78) and nitrocinnamic acids (46) have been prepared in this way. Stoermer and Wehln (177), working with sodium hypochlorite, prepared a number of substituted cinnamic acids from the corresponding unsaturated methyl ketones, and stated that the method was general.

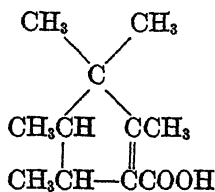
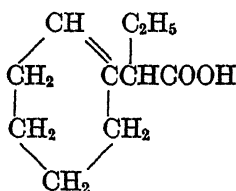
The method is also useful in the degradation of cinnamylideneacetone (40) and its derivatives:



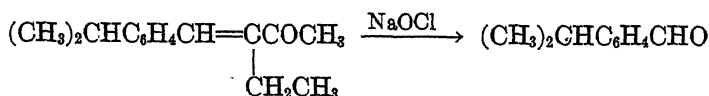
However, in the case of *p*-nitrocinnamylideneacetone, Einhorn and Gehrenbeck (47) found it necessary to use a hot concentrated solution of sodium hypochlorite, and that the yield of acid was small.

In the aliphatic series Doeuvre (43) has prepared  $\alpha$ -bromo- $\beta$ -methylcrotonic acid from the corresponding ketone by the action of potassium hypobromite. Similarly, Cuculescu (34) obtained  $\beta$ -methylcrotonic acid in a 75 per cent yield from mesityl oxide, using iodine in alkali. Likewise, the hexenoic acid,  $(\text{CH}_3)_2\text{C}=\text{CBrCH}_2\text{CH}_2\text{COOH}$ , was made from the corresponding bromomethylheptenone (208).

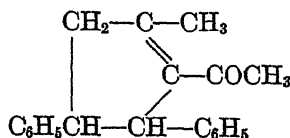
Examples of cyclene acids which have been synthesized by this method are tetrahydrotoluic acid (153),  $\beta$ -campholytic acid (16), and  $\alpha$ -ethyl- $\Delta^1$ -cyclohexenylacetic acid (108).

 $\beta$ -Campholytic acid $\alpha$ -Ethyl- $\Delta^1$ -cyclohexenyl-acetic acid

That the method is not universally applicable to unsaturated methyl ketones is indicated by the experience of Warunis and Lekos (214), who found that the condensation product of cuminaldehyde and methyl propyl ketone was reconverted to the aldehyde by the action of sodium hypochlorite.



Similarly, Harries and Hübner (77) found that 1-acetyl-2-methyl-4,5-diphenyl-1-cyclopentene



resisted degradation by the hypobromite method. It was suggested that this may be due to steric hindrance.

#### D. Aromatic acids

More useful in general is the preparation of aromatic acids by the haloform degradation of the corresponding acetophenones. This method has been widely used and owes its importance to the fact that the introduction of a carboxyl group on an aromatic ring is usually much more difficult than the introduction of the acetyl group. By this method Mills (135) obtained durylic acid from acetylpsedocumene in an 80 per cent yield (171). Van Arendonk and Cupery (191) have worked out the optimum conditions for degrading acetophenones to the corresponding benzoic acids and have obtained yields which vary from 85 to 96 per cent of the theoretical amounts. According to these investigators, hypochlorite is much superior to hypobromite or hypoiodite. They found, however, that hydroxyl or

nitro groups interfere with the reaction.<sup>8</sup> An interesting application of this method is the preparation of 2,2',6,6'-tetramethoxy-3,3'-dicarboxy-diphenyl from the corresponding diacetyl compound (192).

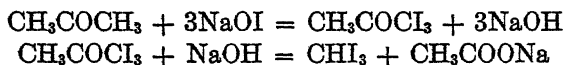
#### VII. THE MECHANISM OF THE HALOFORM REACTION

Since alcohols which give the haloform reaction are capable of being oxidized to methyl ketones or acetaldehyde, it has been assumed that in all these cases either acetaldehyde or a methyl ketone is formed as an intermediate. There is no example of the haloform reaction in which such carbonyl compounds may not be involved. It seems justifiable, then, to limit the discussion of mechanism to compounds which contain the acetyl group united with either hydrogen or carbon.

From the work of Liebig (142), Orndorff and Jessel (143), Zincke (222), Abbott (1) and others it is clear that the generalized equation for the reaction is of the form



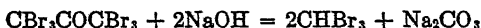
where R may be hydrogen or any group of atoms, provided only that the group contains a carbon atom to which the carbonyl group is joined directly. MOX represents a hypohalite. An inspection of the equation shows that two distinctly different types of processes are involved in the haloform reaction—halogenation and chain cleavage. That the halogenation precedes the cleavage was assumed by Orndorff and Jessel (143), Zincke (222), Noyes (140), and others, who advanced the hypothesis that a trihalomethyl derivative was formed as an intermediate and that this was then cleaved by the alkali. In the case of the formation of iodoform from acetone the equations took the following form:



That this represents the actual sequence is now beyond question,<sup>9</sup> and it is

<sup>8</sup> *m*-Nitroacetophenone, *p*-hydroxyacetophenone, resacetophenone, and nitro-resacetophenone cannot be converted to acids by this method (191). However, since compounds of these types do give iodoform when treated according to the procedure of Fuson and Tullock (62), it is clear that the difficulty in the above cases is not due to the failure of the haloform reaction to take place.

<sup>9</sup> In the case of acetone the hexahalo derivative has also been suggested as an intermediate (222, 159). Since, however, Weidel and Gruber (215) found that hexabromoacetone reacts almost quantitatively with sodium hydroxide according to the equation



it is clear that the hexabromoketone cannot be an intermediate in the main reaction, for sodium acetate is always a principal product.

certain that in the haloform reaction we have to deal with halogenation followed by chain cleavage. It will be convenient to consider the two processes separately.

*A. The halogenation phase of the haloform reaction*

That the trihaloacetyl derivatives actually are intermediates seems certain, inasmuch as it has been possible in certain cases to interrupt the reaction before chain cleavage occurs. This has been accomplished in the case of pinacolone (151) and with a number of methyl aryl ketones carrying

TABLE 1

*Trihalomethyl phenyl ketones prepared by the use of hypohalites*

POSITIONS ON THE PHENYL GROUP						REFERENCE
1	2	3	4	5	6	
COCCl <sub>3</sub>	CH <sub>3</sub>		CH <sub>3</sub>		CH <sub>3</sub>	63
COCCl <sub>3</sub>	CH <sub>3</sub>		CH <sub>3</sub>		CH <sub>3</sub>	63
COCCl <sub>3</sub>	OCH <sub>3</sub>		OCH <sub>3</sub>		OCH <sub>3</sub>	59
COCCl <sub>3</sub>	OCH <sub>3</sub>		OCH <sub>3</sub>		OCH <sub>3</sub>	59
COCCl <sub>3</sub>	Br		Br		Br	61
COCCl <sub>3</sub>	Br		Br		Br	61
COCCl <sub>3</sub>	CH <sub>3</sub>	NO <sub>2</sub>	CH <sub>3</sub>	NO <sub>2</sub>	CH <sub>3</sub>	63
COCCl <sub>3</sub>	CH <sub>3</sub>	NO <sub>2</sub>	CH <sub>3</sub>	NO <sub>2</sub>	CH <sub>3</sub>	63
COCCl <sub>3</sub>	Cl		Cl		Cl	57
COCCl <sub>3</sub>	CH <sub>3</sub>	Br	CH <sub>3</sub>	Br	CH <sub>3</sub>	55
COCCl <sub>3</sub>	CH <sub>3</sub>	Br	CH <sub>3</sub>	Br	CH <sub>3</sub>	55
COCCl <sub>3</sub>	CH <sub>3</sub>	COCCl <sub>3</sub>	CH <sub>3</sub>		CH <sub>3</sub>	70
COCCl <sub>3</sub>	CH <sub>3</sub>	COCCl <sub>3</sub>	CH <sub>3</sub>		CH <sub>3</sub>	70
COCCl <sub>3</sub>	CH <sub>3</sub>	CH <sub>3</sub>		CH <sub>3</sub>	CH <sub>3</sub>	70
COCCl <sub>3</sub>	CH <sub>3</sub>	CH <sub>3</sub>		CH <sub>3</sub>	CH <sub>3</sub>	70
COCCl <sub>3</sub>	CH <sub>3</sub>	CH <sub>3</sub>	CH <sub>3</sub>		CH <sub>3</sub>	70
COCCl <sub>3</sub>	CH <sub>3</sub>	CH <sub>3</sub>	CH <sub>3</sub>		CH <sub>3</sub>	70
COCCl <sub>3</sub>	CH <sub>3</sub>	CH <sub>3</sub>	CH <sub>3</sub>	COCCl <sub>3</sub>	CH <sub>3</sub>	70
COCCl <sub>3</sub>	CH <sub>3</sub>	CH <sub>3</sub>	CH <sub>3</sub>	COCCl <sub>3</sub>	CH <sub>3</sub>	70

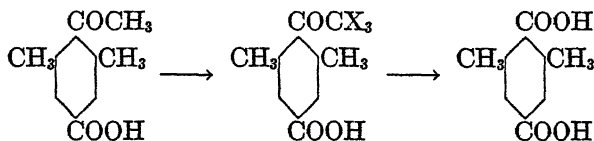
two ortho substituents. Thus, with solutions of hypohalites, acetomesitylene is converted to  $\alpha, \alpha, \alpha$ -trihaloacetomesitylene (63). Table 1 lists the trihalomethyl phenyl ketones which have been prepared by the use of hypohalites.

A study has been made of the influence of various ortho substituents on the stability of the trihalomethyl ketones. As table 1 shows, the substituents which were used in this study are methyl (63, 70), bromine (61), chlorine (57), and methoxyl (59). Of these, the methoxyl group seems to have the least effect in retarding the cleavage phase, although even in this

case the trihalomethyl ketones are sufficiently stable toward alkalis to make it possible to prepare them in alkaline solution.

Without exception, the trihalomethyl ketones mentioned above were insoluble in the alkaline media used in their preparation. Experiments with alcohol and pyridine as solvents indicated, however, that the great stability of these compounds was due in part, at least, to their insolubility in alkali. To demonstrate clearly the effect of solubility on the stability of hindered trihalomethyl ketones, Bull and Fuson synthesized 3,5-dimethyl-4-acetylbenzoic acid and converted it to the corresponding trihalomethyl ketonic acids (24). These compounds were all soluble in alkali, hence the insolubility factor was eliminated.

The synthesis of these compounds in alkaline solutions demonstrates that they are moderately stable toward alkalis. Moreover, when dissolved in sodium hydroxide solutions at 0°C. they could be recovered unchanged. However, if these solutions were allowed to stand, or were warmed, the trihalomethyl ketonic acid underwent rapid cleavage to dimethylterephthalic acid:



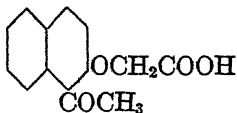
These results indicated that the solubility factor is of considerable importance to the stability of the trihalomethyl ketones in alkali. The ease with which cleavage takes place is greatly increased by the solubility of the ketone in alkali, and, conversely, the stability of trihalomethyl ketones is greatly enhanced by their insolubility in aqueous alkali.

A more striking demonstration of this fact has been furnished by the same authors; they compared the reactions of methyl aryl ketones in which (a) the solubility factor was changed while the steric factor remained constant, and (b) the steric factor was changed while the solubility factor remained constant (25).

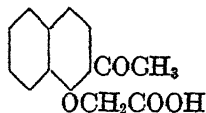
Such a comparison was satisfactorily afforded by the study of the three ketones 1-aceto-2-naphthyl methyl ether (I), 1-aceto-2-naphthoxyacetic acid (II), and 2-aceto-1-naphthoxyacetic acid (III).



I



II



III



In the first two, the solubility factor is changed while the steric factor remains approximately constant, since the molecular volume of the acetoxyl group, while unknown, would certainly not be less than that of the methoxyl group. In the last two the steric factor is changed while the solubility factor remains approximately constant, since the solubility of these position isomers in alkali may be assumed to be the same.

1-Aceto-2-naphthyl methyl ether (I) and the corresponding trihalomethyl derivatives were prepared by Fuson, Farlow, and Stehman (59), who found it necessary to reflux the trihaloacetyl derivatives for four hours with 20 per cent alkali before the cleavage to the corresponding acid was complete.

The hindered ketone (II), when dissolved in solutions of sodium hypohalite, gave 1-trihaloacetyl-2-naphthoxyacetic acid (25). The synthesis of these compounds in alkaline solutions demonstrates that they are moderately stable toward alkalis. Moreover, when dissolved in sodium hydroxide solutions at 0°C., they could be recovered unchanged. Yet the effect of the solubility on the stability of these trihalomethyl ketonic acids was indicated by their conversion into the dibasic acid when warmed with dilute alkali for half an hour, *or in approximately one-eighth the time required for the cleavage of the alkali-insoluble trihalomethyl ketones derived from 1-aceto-2-naphthyl methyl ether (I) (59)*. Hence the cleavage of these two types of trihalomethyl ketones possessing approximately equal steric factors, and differing essentially only in solubility, indicates that the introduction of a solubilizing group leads to an increase in the ease of cleavage, and, conversely, that the insolubility of the trihalomethyl ketones in alkali greatly enhances their apparent stability.

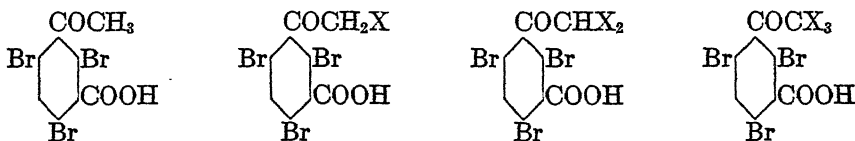
When the unhindered ketone (III) was treated with ice-cold solutions of hypohalites, quantitative cleavage to the dibasic acid resulted (25). This result was observed under conditions identical with those obtaining when the trihalomethyl ketones of the hindered acid were prepared. Even when the reaction time was reduced to five minutes and when, in addition, an attempt was made to salt out any trihalomethyl ketonic acids which might have been formed,<sup>10</sup> cleavage to the dibasic acid was quantitative. In sharp contrast to this behavior is the fact that the hindered trihalomethyl ketonic acids are readily prepared in cold dilute alkaline hypohalite solutions and, indeed, can be recovered unchanged from dilute alkali at room temperature.

It was, of course, highly probable that these trihaloacetyl derivatives were formed by a stepwise halogenation of the corresponding ketones, but

<sup>10</sup> These trihaloacetylbenzoic acids are relatively insoluble in alkali. Indeed, the isolation of the insoluble sodium salt has been used as a convenient means of purifying these derivatives (24).

because of the rapidity of the reaction, this aspect of the halogenation phase had heretofore eluded demonstration. An intermediate monohalogen derivative was isolated by Fuson, Bertetti, and Ross (57), who succeeded in preparing the monochloroacetyl derivative of 2,4,6-trichloroacetophenone in sodium hypochlorite solution.

The isolation of all the intermediate haloacetyl derivatives has recently been accomplished (60) by the treatment of 2,4,6-tribromo-3-acetylbenzoic acid with alkaline hypohalites for various periods of time, ranging from fifteen minutes to twenty-four hours.



The isolation of these compounds clearly demonstrates that the halogenation phase of the haloform reaction involves a step-by-step halogenation to the trihaloacetyl derivatives.

The mechanism of the halogenation reaction still remains obscure. This process is often assumed<sup>11</sup> to proceed through the enol form of the methyl ketone in order to account for the fact that the iodoform and bromoform reactions are pseudo-unimolecular.<sup>12</sup> Bartlett (9) and others have assumed that the rate-controlling reaction is the enolization, and that all subsequent reactions are extremely rapid.

#### *B. The cleavage phase of the haloform reaction*

The evidence for the final step in the haloform reaction was set forth by Liebig, who prepared chloroform by the action of alkali on chloral (126, 127):



Many compounds containing the trihaloacetyl grouping ( $\text{CX}_3\text{CO}$ —) joined to hydrogen or to carbon are now known and, except in certain special cases already noted which involve steric hindrance, these are readily cleaved by alkalis to give a haloform and the salt of an acid:  $\text{RCOCX}_3 + \text{NaOH} = \text{CHX}_3 + \text{RCOONa}$ . Tables 2, 3, and 4 contain lists of such compounds.

<sup>11</sup> See, for example, Pieroni and Tonniolo (146), Bartlett (9), Sukhnevich and Chilingaryan (180).

<sup>12</sup> The reaction of acetone with hypochlorite is a bimolecular reaction, according to Bartlett (9); in this connection it is interesting to note that Böttiger and Kötze (18) have found the alkaline decomposition of chloral hydrate is itself a second-order reaction when a great excess of alkali is present.

Trihaloacetyl derivatives of more complex types have also been prepared and subjected to cleavage with alkali. Among these are derivatives of  $\alpha$ -naphthol (87), tetralin (89),  $\beta$ -naphthyl methyl ether (59),  $\beta$ -naphthoxyacetic acid (25), and pyrrole (54).

In most cases strong alkalies such as sodium, potassium, calcium, and barium hydroxides are used to bring about the cleavage, but sodium car-

TABLE 2  
*Trihalomethyl aldehydes and ketones (aliphatic)*

COMPOUND	REFERENCE	COMPOUND	REFERENCE
$\text{CCl}_3\text{CHO}$ .....	127, 149, 216, 49	$\text{CCl}_3\text{COCH}_2\text{Br}$ .....	29
$\text{CBr}_3\text{CHO}$ .....	130	$\text{CBr}_3\text{COCBr}_3$ .....	215, 75
$\text{CCl}_3\text{COCH}_3$ .....	111, 136, 29	$\text{CCl}_2\text{BrCOCBr}_2$ .....	76
$\text{CBr}_3\text{COCH}_3$ .....	50	$\text{CBr}_3\text{COCBr}_2$ .....	76
$(\text{CH}_3)_2\text{CHCOCBr}_3$ .....	97	$\text{CBr}_3\text{COCHBr}_2$ .....	28
$\text{CCl}_3\text{COCH}_2\text{Cl}$ .....	29	$\text{CBr}_3\text{COCOCBr}_3$ .....	95
$\text{CCl}_3\text{COCHCl}_2$ .....	29	$\text{CCl}_3\text{COCBr}_2$ .....	221
$\text{CCl}_3\text{COCBr}_3$ .....	29	$\text{CCl}_3\text{COCBr}_2$ .....	225
$\text{CBr}_3\text{COCH}_2\text{Cl}$ .....	29	$(\text{CH}_3)_2\text{CCOCBr}_3$ .....	151

TABLE 3  
*Trihalomethyl ketonic acids (aliphatic)*

COMPOUND	REFERENCE
$\text{CBr}_3\text{COCH}=\text{CHCOOH}$ .....	217
$\text{CBr}_3\text{COCBr}=\text{CHCOOH}$ (?).....	217
$\text{CCl}_3\text{COCH}=\text{CHCOOH}$ .....	100
$\text{CCl}_3\text{COCBr}=\text{CHCOOH}$ .....	221, 226
$\text{CCl}_3\text{COCHBrCHBrCOOH}$ .....	100
$\text{CCl}_3\text{COCBr}=\text{CHCOOH}$ .....	231, 227
$\text{CCl}_3\text{COCBr}=\text{C}(\text{CH}_3)\text{COOH}$ .....	176, 100, 221, 220
$\text{CCl}_3\text{COCBr}=\text{CHCCl}_2\text{COOH}$ .....	221
$\text{CCl}_3\text{COCBr}=\text{CClCCl}_2\text{COOH}$ .....	225
$\text{CCl}_3\text{CO}(\text{CH}_2)_8\text{COOH}$ .....	137
$\text{CCl}_3\text{COCBr}_2\text{CHClCCl}_2\text{COOH}$ .....	231

bonate (217) and ammonium hydroxide (29) have been employed also. That sodium carbonate is not always sufficient is shown by the work of Zincke and Egly (223), who carried out the following reaction in *hot* sodium carbonate solution:

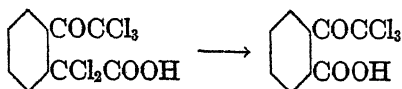


TABLE 4

*Trihalomethyl aryl ketones and ketonic acids*

POSITIONS ON THE PHENYL GROUP						REF- ER- ENCE
1	2	3	4	5	6	
COCl <sub>3</sub>						31
COCCl <sub>3</sub>						{ 65
COCCl <sub>3</sub>						{ 41
COCCl <sub>3</sub>						137
COCCl <sub>3</sub>			CCl <sub>3</sub>			52
COCCl <sub>3</sub>	CCl <sub>3</sub>		CCl <sub>3</sub>			51
COCCl <sub>3</sub>	COOR		CCl <sub>3</sub>			51
COCCl <sub>3</sub>	COOR			CCl <sub>3</sub>		51
COCCl <sub>3</sub>	CCl <sub>3</sub>		CCl <sub>3</sub>	COCCl <sub>3</sub>		51
COCCl <sub>3</sub>	CH <sub>3</sub>	COCCl <sub>3</sub>	CCl <sub>3</sub>		CCl <sub>3</sub>	51
COCCl <sub>3</sub>	CCl <sub>3</sub>			CCl <sub>3</sub>		51
COCCl <sub>3</sub>			OCH <sub>2</sub> CH <sub>3</sub>			87
COCCl <sub>3</sub>		SO <sub>3</sub> H	OCH <sub>2</sub> CH <sub>3</sub>			87
COCCl <sub>3</sub>			OCH <sub>3</sub>			87
COCCl <sub>3</sub>	CH <sub>3</sub>		OCH <sub>3</sub>			87
COCCl <sub>3</sub>		CH <sub>3</sub>	OCH <sub>2</sub> CH <sub>3</sub>			87
COCCl <sub>3</sub>	OCH <sub>3</sub>			CH <sub>3</sub>		87
COCCl <sub>3</sub>		OCH <sub>3</sub>	OCH <sub>3</sub>			87
COCCl <sub>3</sub>		Br	OCH <sub>3</sub>			87
COCCl <sub>3</sub>			OC <sub>6</sub> H <sub>5</sub>			87
COCCl <sub>3</sub>	CH <sub>3</sub>		OH	CH <sub>3</sub>		89
COCCl <sub>3</sub>	CH <sub>3</sub>		OH	CH(CH <sub>3</sub> ) <sub>2</sub>		89
COCCl <sub>3</sub>		CH <sub>3</sub>	OH		CH(CH <sub>3</sub> ) <sub>2</sub>	89
COCCl <sub>3</sub>		CH <sub>3</sub>	OH			88
COCCl <sub>3</sub>			OH			88
COCCl <sub>3</sub>	OH		CH <sub>3</sub>			88
COCCl <sub>3</sub>	CH <sub>3</sub>		OH			88
COCCl <sub>3</sub>	OCH <sub>3</sub>			CH <sub>3</sub>		88
COCCl <sub>3</sub>	OH			CH <sub>3</sub>		88
COCCl <sub>3</sub>	COOH		COOH			51
COCCl <sub>3</sub>	COOH			COOH		51
COCCl <sub>3</sub>	Cl	Cl	Cl	Cl	COOH	230
COCCl <sub>3</sub>			COOH			52
COCCl <sub>3</sub>	COOH					229
COCCl <sub>3</sub>			COOCH <sub>3</sub>			52
COCCl <sub>3</sub> Br	COOH					229
COCCl <sub>3</sub> Br <sub>2</sub>	COOH					229
COCCl <sub>3</sub>	COOH		COOH		Br	224
COCCl <sub>3</sub>	CCl <sub>3</sub> COOH					223
COCCl <sub>3</sub>						{ 229
COCCl <sub>3</sub>	COOH					{ 64

However, the resulting ketonic acid was decomposed by alkali into chloroform and phthalic acid.

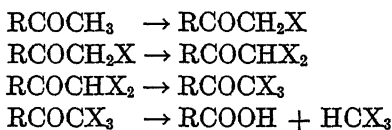
An entirely new light was thrown on the nature of the cleavage of trihalomethyl ketones by an experiment of Jackson and Adams (95). Working with hexabromodiacetyl, they found that cleavage was induced by sodium acetate:



Houben and Fischer have subsequently demonstrated that the process is catalytic. In the presence of sodium as a catalyst, trichloromethyl ketones react with alcohols to give chloroform and esters:  $\text{RCOCCl}_3 + \text{ROH} = \text{RCOOR} + \text{CHCl}_3$  (90). Even more significant is the discovery that trichloroacetophenone, although unchanged by seven hours' heating with water at  $170^\circ\text{C}$ ., is immediately decomposed by the addition of one drop of potassium hydroxide solution to a methyl alcohol solution of the ketone (91). That the cleavage is catalytic is clearly proved by the fact that the ketone is smoothly decomposed by being heated for several hours with water containing a little potassium acetate:



We may conclude with certainty that the mechanism of the haloform reaction involves a stepwise halogenation of compounds which undergo the reaction, followed by chain cleavage of the trihaloacetyl derivatives thus formed, according to the following scheme:



#### VIII. REACTIONS RELATED TO THE HALOFORM REACTION

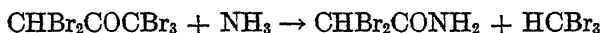
From what has gone before, it is evident that the haloform reaction involves several types of processes—halogenation, chain cleavage, and, if we include alcohols and amines, oxidation. Moreover, each of these represents a very large group of reactions of which the one under consideration is but a single type. It is beyond the scope of the present article to attempt a survey of the work which has been done on reactions which are related to the haloform reaction, yet there are a few degradation processes which resemble the haloform reaction so closely as to warrant mention in this connection. Those mentioned below not only have a formal similarity to the haloform reaction, but resemble it also in the great ease with which they proceed.

*A. Modifications of the haloform reaction*

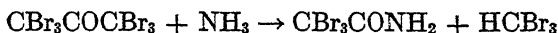
One of the most interesting modifications of the haloform reaction was reported by Chattaway and Baxter (27), who found that nitrogen triiodide reacted with acetophenone to give benzamide and iodoform. The triiodoacetophenone was postulated as an intermediate:



The cleavage in this case appears to be produced by ammonia. Amide formation by the action of ammonia was reported as early as 1886 by Cloëz (29), who carried out the following reaction

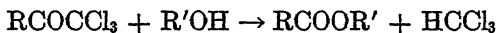


and many others of the same type. Similarly, Weidel and Gruber (215) obtained tribromoacetamide from hexabromoacetone by the action of ammonia.



Similar results were obtained by Hantzsch (76) with other hexahaloacetones. Likewise, ammonia and aniline were found to act upon hexabromodiacetyl to give, respectively, oxamide and oxanilide (95). Cloëz (29) states that similar decompositions can be accomplished by the action of toluidine, allylamine, diethylamine, ethylene diamine, and urea on trihaloacetyl compounds. A result of especial interest has been reported by Datta and Prosad (36) who, by using iodine and ammonia, obtained iodoform from di- and tri-ethylamine.

As has already been mentioned, Houben and Fischer discovered a modification of the haloform reaction in which the trihaloacetyl compound is decomposed by alcohol in the presence of a small amount of sodium (90). Esters and a haloform are formed in this reaction according to the equation:

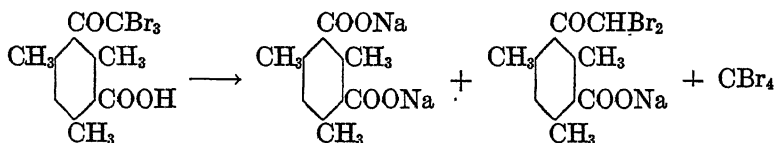
*B. Reactions analogous to the haloform reaction*

It is a well-known fact that hypohalites react with compounds containing active hydrogen atoms in such a manner as to replace the latter by halogen atoms. In this review, however, we shall mention only examples of this type of reaction in which *hydrogen on carbon* is replaced. Examples are known of the replacement of the hydrogen atoms of methylene (58, 56) and methine (114) groups by halogen by the action of hypohalites. It is especially noteworthy that in these types of halogenation the methylene and the methine groups are not necessarily adjacent to a carbonyl group.

Of the examples in which a carbonyl group is not involved, the most striking are the acetylenes (15), cyclopentadiene (178), and the haloforms (198, 32, 8). As has already been stated, the conversion of a haloform to the corresponding tetrahalomethane is a general reaction and takes place quantitatively (83, 73, 38, 8).

From these facts emerges the generalization that mobile or active hydrogen atoms on carbon are replaced by halogen in the treatment with hypohalites. It has been suggested (146) that this process involves enolization, followed by addition of the hypohalite to the enol form. From the foregoing examples, however, it is clear that replacement of hydrogen takes place in some compounds which are incapable of enolization. Moreover, no proof has yet been found that an enol form is involved in any of these cases. Further, it is worth noting that a halogen atom which has replaced an active hydrogen atom has certain peculiar properties, such as that of being itself a halogenating agent, *and this property is shared by the carbon tetrahalides in which enolization to a hypobromite structure cannot be assumed to account for the unusual reactivity of the atom involved* (93).

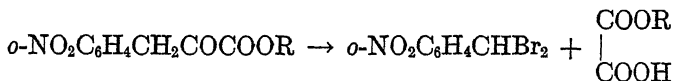
An interesting example of this peculiar behavior was discovered by Johnson and Fuson (98), who found that 2,4,6-trimethyl-4-tribromoacetylbenzoic acid was converted by alkali into the salts of 2,4,6-trimethylisophthalic acid and 2,4,6-trimethyl-4-dibromoacetylbenzoic acid, according to the following scheme:



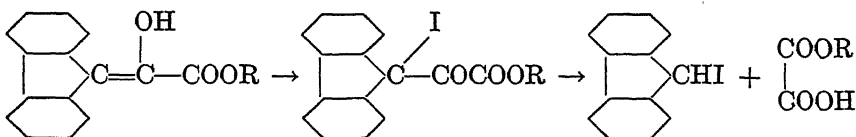
Carbon tetrabromide was also formed in the reaction, and it must be supposed that it results from the brominating action of the tribromomethyl ketone on bromoform.

An interesting extension of the haloform reaction is seen in the formation of iodoform by the action of hypoiodite on silver, copper, and mercury acetylides (118). Here the results suggest hydration to give a methyl ketone, which thereafter undergoes the normal haloform reaction. Recent work on this interesting reaction indicates that this behavior of acetylides may be general (137).

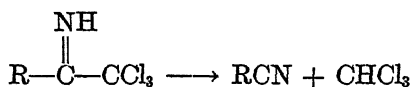
Another reaction, clearly of the haloform type, was observed by Reissert (150), who found that *o*-nitrophenylpyruvic ester is converted by sodium hypobromite into *o*-nitrobenzal bromide, a compound which may be regarded as bromoform in which one atom of bromine is replaced by an aryl group.



A similar case was discovered by Kuhn and Levy (114), in which a bi-phenylene radical replaces two of the halogen atoms of the haloform. These investigators obtained 9-iodofluorene by the action of hypoiodite on  $\alpha$ -hydroxy- $\beta$ -diphenyleneacrylic ester:

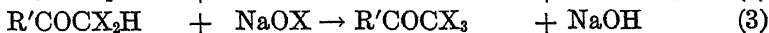


Another interesting and synthetically important variation of the haloform degradation is the catalytic decomposition which occurs when solutions of imides of trichloromethyl ketones are allowed to stand over alkalis. Chloroform and the corresponding nitrile are produced.

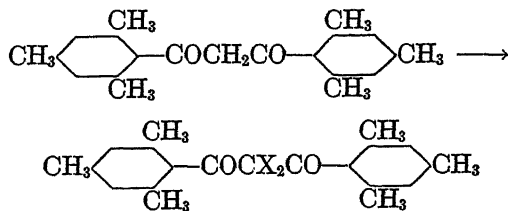


A number of examples of this type have been found by Houben and Fischer, who state that the reaction is general (92).

The action of hypohalites on  $\beta$ -diketones has already been discussed; it gives rise to the formation of a haloform or of a closely related type of compound. Thus, acetyl-*p*-bromoacetophenone (218), dihydroresorcinol (196), methone (196, 106), and bromomethone (106) yield bromoform with hypobromite. The process has been formulated as follows (218):



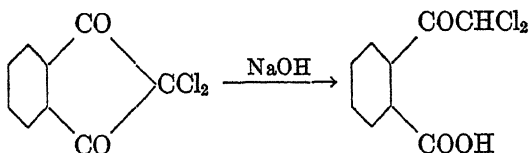
The first step is exemplified by the case of di( $\beta$ -isoduryloyl)methane, which is transformed by hypohalites into the corresponding dihalo-diketone (218):



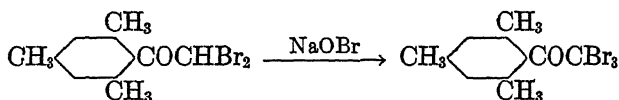


Here the hindrance offered by the mesityl groups arrests the reaction at the first stage.

The second step—the cleavage of the dihalodiacylmethanes by alkalis—is illustrated by the reaction of dichlorodiketohydrindene with alkali to give *o*-dichloroacetylbenzoic acid (228):

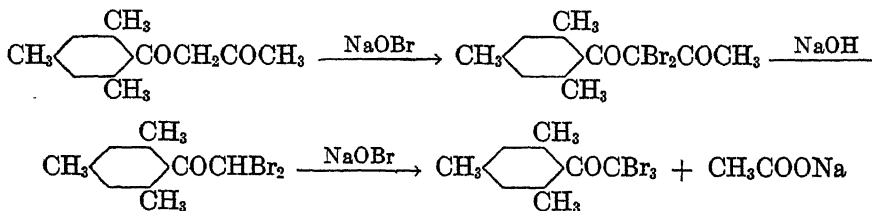


The third step is illustrated by the reaction of  $\alpha,\alpha$ -dibromoacetomesitylene with hypobromite, the corresponding trihaloacetyl derivative being produced (56):



The last step—the cleavage of trihaloacetyl compounds—is, of course, illustrated by a great number of known compounds.<sup>13</sup>

A combination of the first three steps is illustrated in the reaction of  $\alpha$ -acetylacetomesitylene, which is converted by alkaline hypobromite to  $\alpha,\alpha,\alpha$ -tribromoacetomesitylene (218):



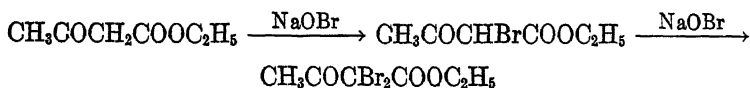
It is interesting to note that this is also an example of a methyl ketone which does not give the haloform reaction; the acetyl group is eliminated in the form of acetic acid.

Closely similar to the behavior of  $\beta$ -diketones is that of  $\beta$ -ketonic esters. In the case of acetoacetic ester the product with calcium hypochlorite is dichloroacetic acid, formed as follows (94):



<sup>13</sup> See tables 2, 3, and 4.

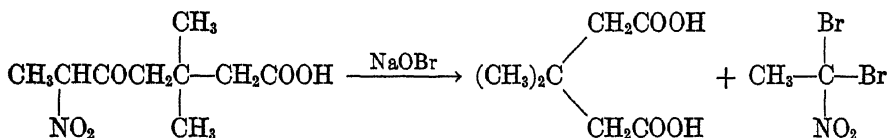
That this correctly represents the mechanism is supported by Brühl's observation (23) that bromine in alkali converts ethyl acetoacetate to ethyl  $\alpha$ -bromoacetoacetate and ethyl  $\alpha, \alpha$ -dibromoacetoacetate:



The behavior of methyl ketones is paralleled by that of nitromethyl ketones. Thiele and Haeckel (182) found that  $\alpha$ -nitroacetophenone gives dibromonitromethane and bromopicrin with hypobromite:

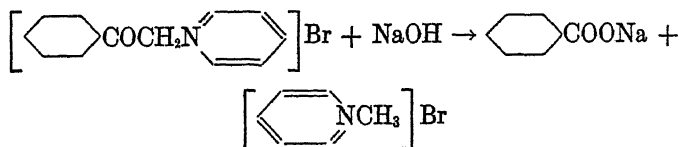


A similar degradation was observed by Toivonen, Osara, and Ollila (189):



It is interesting to note at this point that methyl sulfones and sulfoxides do not give haloforms with hypohalites (169).

One of the outstanding characteristics of the haloform type of reaction is the ease with which it takes place. In this respect it is set off sharply from the cleavage by alkali of ordinary ketones,  $\beta$ -diketones,  $\beta$ -ketonic esters, and the like. A reaction very similar to the haloform reaction in this respect is the cleavage of keto quaternary ammonium salts of the type illustrated by phenacylpyridinium bromide (6). Schmidt and van Ark (160) have shown that the reaction takes place according to the equation:



Kröhnke (110), Krollpfeiffer and Müller (113), and others (4) have confirmed and extended these results.<sup>14</sup> It appears that this cleavage of  $\beta$ -keto pyridinium halides is a general reaction of potential synthetic importance.

<sup>14</sup> Auwers and Lämmerhirt (3) showed that certain  $\alpha$ -halogen ketones could be degraded to acids by treatment with pyridine; it seems probable that here also the pyridinium halide is formed as an intermediate.

## IX. SUMMARY

1. Methyl ketones and acetaldehyde yield haloforms when treated with hypohalites. The process involves stepwise halogenation followed by chain cleavage of the resulting trihalomethyl derivative.

2. A few methyl ketones fail to react normally. This is sometimes due to the presence of a group which is very sensitive to the reagent and which diverts the reaction from the normal course. More frequently, however, the interference is steric. If the acetyl group is joined to a radical which is highly branched the reaction may be interrupted at the cleavage phase, the product being a trihalomethyl ketone.

3. The haloform reaction occurs also with certain alcohols, amines,  $\beta$ -diketones, olefins, oximes, esters, and acetylides. In all of these cases, however, the compounds are convertible to methyl ketones (or acetaldehyde) by processes such as oxidation, hydrolysis, and halogenation,—processes which are known to be brought about by hypohalite solutions. In other words, it would appear that all of the compounds which undergo the haloform reaction do so by first being transformed into methyl ketones (or acetaldehyde).

4. Considerable differences have been observed in the results obtained with the various hypohalites. Thus, hypochlorite and hypobromite often attack compounds which are unaffected by hypoiodite. In particular, the triiodomethyl derivatives, postulated as intermediates, have never been prepared. Even in those cases in which the corresponding trichloro and tribromo compounds are readily obtainable the iodination does not go to completion but yields only diiodomethyl derivatives.

The authors are indebted to Professors Mario Passerini and Raoul Poggi of the Royal University of Florence, Italy, for permission to use the libraries of that institution.

## X. REFERENCES

- (1) ABBOTT: *J. Phys. Chem.* **7**, 84 (1903).
- (2) ARACHEQUESNE: *Compt. rend.* **110**, 642 (1890).
- (3) AUWERS AND LÄMMERHIRT: *Ber.* **53**, 433 (1920).
- (4) BABCOCK, NAKAMURA, AND FUSON: *J. Am. Chem. Soc.* **54**, 4407 (1932).  
BABCOCK AND FUSON: *J. Am. Chem. Soc.* **55**, 2946 (1933).
- (5) BAEYER: *Ber.* **29**, 25 (1896).
- (6) BAMBERGER: *Ber.* **20**, 3344 (1887).
- (7) BARBIER AND GRIGNARD: *Compt. rend.* **147**, 597 (1908).
- (8) BARTAL: *Chem.-Ztg.* **29**, 377 (1909).
- (9) BARTLETT: *J. Am. Chem. Soc.* **56**, 967 (1934).
- (10) BATES, MULLALLY, AND HARTLEY: *J. Chem. Soc.* **123**, 401 (1923).
- (11) BELOHOUBEK: *Ann.* **165**, 350 (1873).
- (12) BENESCH: *Chem.-Ztg.* **50**, 98 (1926).

- (13) BENTLEY AND PERKIN: *J. Chem. Soc.* **73**, 52 (1898).
- (14) BERGMANN, ZERVAS, AND ENGLER: *Ann.* **508**, 25 (1933).
- (15) BILTZ AND KÜPPERS: *Ber.* **37**, 4412 (1904).
- (16) BLANC: *Bull. soc. chim.* [4] **5**, 24 (1909).
- (17) BÖSEKEN: *Rec. trav. chim.* **29**, 99 (1910).
- (18) BÖTTIGER AND KÖTZ: *J. prakt. Chem.* [2] **65**, 481 (1902).
- (19) BONNET: *J. prakt. Chem.* **10**, 207 (1837).
- (20) BORSCHKE: *Ber.* **44**, 2595 (1911).
- (21) BOUCHARDAT: *J. Pharm. Chem.* **23**, 1 (1837); *Ann.* **22**, 225 (1837).
- (22) BOUVEAULT AND LEVALLOIS: *Compt. rend.* **146**, 180 (1908).
- (23) BRÜHL: *Ber.* **36**, 1731 (1903).
- (24) BULL AND FUSON: *J. Am. Chem. Soc.* **55**, 3424 (1933).
- (25) BULL AND FUSON: *J. Am. Chem. Soc.* **56**, 736 (1934).
- (26) CAHOURS: *Ann. chim.* [3] **19**, 488 (1847).
- (27) CHATTAWAY AND BAXTER: *J. Chem. Soc.* **103**, 1986 (1913).
- (28) CLOËZ: *Ann.* **122**, 121 (1862).
- (29) CLOËZ: *Ann. chim.* [6] **9**, 145 (1886).
- (30) COHEN: *Pharm. Weekblad* **69**, 363 (1932).
- (31) COLLET: *Bull. soc. chim.* [3] **23**, 831 (1900).
- (32) COLLIE: *J. Chem. Soc.* **65**, 262 (1894).
- (33) COLLISCHONN: *Z. anal. Chem.* **29**, 562 (1890).
- (34) CUCULESCU: *Bul. Fac. Stiinte Cernauti* **1**, 53 (1927).
- (35) CUCULESCU: *Bul. Fac. Stiinte Cernauti* **2**, 137 (1928).
- (36) DATTA AND PROSAD: *J. Am. Chem. Soc.* **39**, 441 (1917).
- (37) DAVIS: *Archeion* **13**, 11 (1931).
- (38) DEHN: *J. Am. Chem. Soc.* **31**, 1227 (1909).
- (39) DELANGE: *Bull. soc. chim.* [4] **3**, 914 (1908). Cf. DENIGÈS: *Bull. soc. chim.* [3] **29**, 598 (1903).
- (40) DIEHL AND EINHORN: *Ber.* **18**, 2324 (1885).
- (41) DINESMANN: *Compt. rend.* **141**, 202 (1905).
- (42) DOBBIN: *Chemist and Druggist* **115**, 761 (1931).
- (43) DOEUVRE: *Bull. soc. chim.* [4] **39**, 1594 (1926).
- (44) DUMAS: *Ann. chim.* **56**, 113, 122 (1834).
- (45) DUMAS AND PELIGOT: *Ann.* **15**, 9 (1835).
- (46) Dyeworks, Höchst-on-the-Main: *D. R. P.* 21162; *Friedländer* **1**, 28 (1882).  
See also EINHORN AND GRABFIELD: *Ann.* **243**, 363 (1888).
- (47) EINHORN AND GEHRENBECK: *Ann.* **253**, 357 (1889).
- (48) ELLIOTT AND DALTON: *Analyst* **44**, 132 (1919).
- (49) ENKLAAR: *Rec. trav. chim.* **23**, 419 (1904); **24**, 419 (1905); **25**, 297 (1906); **29**, 173 (1910).
- (50) ÉTARD: *Compt. rend.* **114**, 753 (1892).
- (51) FEIST: *Ann.* **496**, 99 (1932).
- (52) FEIST: *Ber.* **67B**, 938 (1934).
- (53) FEYER: *Z. Elektrochem.* **25**, 115 (1919).
- (54) FISCHER AND VIAUD: *Ber.* **64B**, 193 (1931).
- (55) FISHER: *J. Am. Chem. Soc.* **55**, 4241 (1933).
- (56) FISHER, SNYDER, AND FUSON: *J. Am. Chem. Soc.* **54**, 3665 (1932).
- (57) FUSON, BERTETTI, AND ROSS: *J. Am. Chem. Soc.* **54**, 4380 (1932).
- (58) FUSON AND BEVERIDGE: *J. Am. Chem. Soc.* **53**, 1985 (1931).
- (59) FUSON, FARLOW, AND STEHMAN: *J. Am. Chem. Soc.* **53**, 4097 (1931).
- (60) FUSON, JOHNSON, AND BULL: Unpublished work.

- (61) FUSON, LEWIS, AND DU PUIS: *J. Am. Chem. Soc.* **54**, 1114 (1932).
- (62) FUSON AND TULLOCK: *J. Am. Chem. Soc.* **56**, 1638 (1934).
- (63) FUSON AND WALKER: *J. Am. Chem. Soc.* **52**, 3269 (1930).
- (64) GABRIEL AND MICHAEL: *Ber.* **10**, 1551 (1877).
- (65) GAUTIER: *Ann. chim.* [6] **14**, 396 (1888).
- (66) GEELMUYDEN: *Z. anal. Chem.* **35**, 503 (1896).
- (67) GILMAN: *Organic Syntheses*, Coll. Vol. I, p. 512. John Wiley and Sons, New York (1932).
- (68) GLASSTONE: *Ind. Chemist* **7**, 315 (1931).
- (69) GOODWIN: *J. Am. Chem. Soc.* **42**, 39 (1920).
- (70) GRAY, WALKER, AND FUSON: *J. Am. Chem. Soc.* **53**, 3494 (1931).
- (71) GRODZKI: *Ber.* **16**, 512 (1883).
- (72) GUTHRIE: *Am. J. Sci. and Arts* **21**, 64 (1832). See also reference 37.
- (73) HABERMANN: *Ann.* **167**, 174 (1873).
- (74) HALLER AND LASSIEUR: *Compt. rend.* **150**, 1015 (1910).
- (75) HANTZSCH: *Ber.* **20**, 2040 (1887); **21**, 244 (1888).
- (76) HANTZSCH: *Ber.* **22**, 1238 (1889).
- (77) HARRIES AND HÜBNER: *Ann.* **296**, 301 (1897).
- (78) HARRIES AND WARUNIS: *Ann.* **330**, 264 (1904).
- (79) HATCHER AND MUELLER: *Trans. Roy. Soc. Can.* [3] **23**, Sect. 3, 35 (1929).
- (80) HEIKEL: *Chem.-Ztg.* **32**, 75 (1908).
- (81) HENRY AND PAGET: *J. Chem. Soc.* **130**, 72 (1928).
- (82) HINTZ: *Z. anal. Chem.* **27**, 182 (1888).
- (83) HLASIWETZ AND HABERMANN: *Ann.* **159**, 322 (1871).
- (84) HOFFMAN: *J. Am. Chem. Soc.* **51**, 2542 (1929).
- (85) HOFMEISTER: *Z. anal. Chem.* **29**, 632 (1890).
- (86) HOMEYER, WHITMORE, AND WALLINGFORD: *J. Am. Chem. Soc.* **55**, 4209 (1933).
- (87) HOUBEN AND FISCHER: *Ber.* **60B**, 1759 (1927).
- (88) HOUBEN AND FISCHER: *J. prakt. Chem.* [2] **123**, 262 (1929).
- (89) HOUBEN AND FISCHER: *Ber.* **63B**, 2455 (1930).
- (90) HOUBEN AND FISCHER: *Ber.* **64B**, 240 (1931).
- (91) HOUBEN AND FISCHER: *Ber.* **64B**, 2636 (1931).
- (92) HOUBEN AND FISCHER: *Ber.* **66B**, 339 (1933).
- (93) HUNTER AND EDGAR: *J. Am. Chem. Soc.* **54**, 2025 (1932).
- (94) HURD AND THOMAS: *J. Am. Chem. Soc.* **55**, 1646 (1933).
- (95) JACKSON AND ADAMS: *J. Am. Chem. Soc.* **37**, 2529 (1915).
- (96) JERUSALEM: *Biochem. Z.* **12**, 361 (1908).
- (97) JOCIZ: *J. Russ. Phys.-Chem. Soc.* **29**, 104 (1897).
- (98) JOHNSON AND FUSON: *J. Am. Chem. Soc.* **56**, 1417 (1934).
- (99) KEBLER: *J. Am. Chem. Soc.* **19**, 316 (1897).
- (100) KEKULÉ AND STRECKER: *Ann.* **223**, 170 (1884).
- (101) KEPPELER: *Z. angew. Chem.* **18**, 464 (1905).
- (102) KERP AND MÜLLER: *Ann.* **299**, 224 (1898).
- (103) KERSCHBAUM: *Ber.* **33**, 891 (1900).
- (104) KLAR: *Z. anal. Chem.* **35**, 595 (1896).
- (105) KOLTHOFF: *Pharm. Weekbald* **62**, 652 (1925).
- (106) KOMPPA: *Ber.* **32**, 1423 (1899); *Ann.* **368**, 135 (1909). Cf. MAYURANATHAN AND GUHA: *J. Indian Inst. Sci.* **15A**, 131 (1932).
- (107) KOMPPA AND ROSCHIER: *Ann.* **470**, 151 (1929).
- (108) KON AND NARAYANAN: *J. Chem. Soc.* **1927**, 1536.
- (109) KONDO, NOTO, AND TANAKA: *J. Pharm. Soc. Japan* **48**, 1163 (1928).

- (110) KRÖHNKE: Ber. **66B**, 604, 1386 (1933).
- (111) KRAEMER: Ber. **7**, 252 (1874).
- (112) KRÄMER: Ber. **13**, 1000 (1880).
- (113) KROLLPFEIFFER AND MÜLLER: Ber. **66B**, 739 (1933).
- (114) KUHN AND LEVY: Ber. **61B**, 2240 (1928).
- (115) KUNZ: Z. anal. Chem. **59**, 302 (1920).
- (116) LAPWORTH: J. Chem. Soc. **85**, 1222 (1904).
- (117) LAZZARINI: Industria chimica **4**, 771 (1929).
- (118) LE COMTE: J. pharm. chim. [6] **16**, 297 (1902).
- (119) VAN DER LEE: Chem. Weekblad **23**, 444 (1926).
- (120) LEERS: Bull. soc. chim. [4] **39**, 651 (1926).
- (121) LEFORT: Compt. rend. **23**, 229 (1846).
- (122) LESER: Bull. soc. chim. [3] **21**, 548 (1899).
- (123) LESER: Compt. rend. **141**, 1033 (1905).
- (124) LIEBEN: Ann. Spl. Bd. **7**, 218 (1870).
- (125) LIEBEN: Ann. Spl. Bd. **7**, 377 (1870).
- (126) LIEBIG: Ann. Physik Chem. **23**, 444 (1831).
- (127) LIEBIG: Ann. **1**, 182 (1832).
- (128) LIEBIG: Ann. **162**, 161 (1872).
- (129) LOCQUIN: Compt. rend. **153**, 285 (1911).
- (130) LÖWIG: Ann. **3**, 296 (1832).
- (131) MARRIOTT: J. Biol. Chem. **16**, 281 (1913).
- (132) MEERWEIN: Ann. **419**, 158 (1919).
- (133) MESSINGER: Ber. **21**, 3366 (1888).
- (134) MILLON: Compt. rend. **21**, 828 (1845).
- (135) MILLS: J. Chem. Soc. **101**, 2191 (1912).
- (136) MORAWSKY: J. prakt. Chem. **12**, 388 (1875).
- (137) MYDDLETON, BARRETT, AND SEAGER: J. Am. Chem. Soc. **52**, 4405 (1930).
- (138) NENITZESCU AND CANTUNIARI: Ber. **65B**, 807 (1932).
- (139) NENITZESCU AND IONESCU: Ann. **491**, 207 (1931).
- (140) NOYES: Ber. **57B**, 1238 (1924).
- (141) NYBERGH: Ber. **55B**, 1960 (1922).
- (142) OCHI: J. Chem. Ind. Japan **24**, 209 (1921).
- (143) ORNDORFF AND JESSEL: Am. Chem. J. **10**, 363 (1888).
- (144) OWEN AND SIMONSEN: J. Chem. Soc. **134**, 1424 (1932).
- (145) PAULY, GILMOUR, AND WILL: Ann. **403**, 123 (1914).
- (146) PIERONI AND TONNIOLI: Gazz. chim. ital. **43**, II, 620 (1913).
- (147) POGGI: Atti soc. ital. progresso sci., XXI Riunione, **2**, 376 (1933).
- (148) RAKSHIT: Analyst **41**, 246 (1916).
- (149) REICHER: Rec. trav. chim. **4**, 347 (1885).
- (150) REISSERT: Ber. **30**, 1043 (1897).
- (151) RICHARD AND LANGLAIS: Bull. soc. chim. [4] **7**, 464 (1910).
- (152) ROBINEAU AND ROLLIN: Mon. sci. [4] **7**, Pt. I, 272 (1893).
- (153) RUPE AND KUENZY: Helv. Chim. Acta **14**, 701 (1931).
- (154) RUZICKA: Helv. Chim. Acta **2**, 361 (1919).
- (155) RUZICKA AND VAN VEEN: Ann. **468**, 133 (1929).
- (156) RUZICKA AND VAN VEEN: Ann. **468**, 143 (1929).
- (157) SAVARÈ: Gazz. chim. ital. **36**, II, 345 (1906).
- (158) SCHLAGDENHAUFFEN: J. Pharm. [3] **36**, 199 (1859).
- (159) SCHLENK AND BERGMANN: Ausführliches Lehrbuch der organischen Chemie, Vol. 1, p. 109. Franz Deuticke, Leipsic and Vienna (1932).

- (160) SCHMIDT AND VAN ARK: Arch. Pharm. **338**, 321 (1900).
- (161) SCHOORL: Pharm. Weekblad **58**, 209 (1921).
- (162) SEMMLER: Ber. **33**, 273 (1900).
- (163) SEMMLER AND BARTELT: Ber. **41**, 386 (1908).
- (164) SEMMLER AND MCKENZIE: Ber. **39**, 1164 (1906).
- (165) SEMMLER AND RIMPEL: Ber. **39**, 2585 (1906).
- (166) SEMMLER AND VON SCHILLER: Ber. **60**, 1591 (1927).
- (167) SERULLAS: Ann. chim. **20**, 165 (1822); **22**, 172, 222 (1823); **25**, 311 (1824).
- (168) SHINODA AND SATO: J. Pharm. Soc. Japan, No. **548**, 860 (1927).
- (169) SHRINER, UPTON, AND ROLAND: Private communication.
- (170) SIMONSEN: J. Chem. Soc. **121**, 2292 (1922).
- (171) SMITH AND BYRKIT: J. Am. Chem. Soc. **55**, 4305 (1933).
- (172) SOHL AND SHRINER: J. Am. Chem. Soc. **55**, 3828 (1933).
- (173) SOUBEIRAN: Ann. chim. **48**, 131 (1831).
- (174) SPETER: Chem.-Ztg. **55**, 781 (1931); Pharm. Zentralhalle **72**, 628 (1931).
- (175) SQUIBB: J. Am. Chem. Soc. **18**, 1068 (1896).
- (176) STENHOUSE: Ber. **6**, 575 (1873).
- (177) STOERMER AND WEHLN: Ber. **35**, 3549 (1902).
- (178) STRAUS, KOLLEK, AND HEYN: Ber. **63B**, 1868 (1930).
- (179) STRAUS AND KÜHNEL: Ber. **66B**, 1834 (1933).
- (180) SUKHNEVICH AND CHILINGARYAN: J. Gen. Chem. **2**, 783 (1932).
- (181) TEEPLE: J. Am. Chem. Soc. **26**, 170 (1904).
- (182) THIELE AND HAECKEL: Ann. **325**, 11 (1902).
- (183) THOMS: Ber. deut. pharm. Ges. **11**, 3 (1901).
- (184) TIEMANN: Ber. **28**, 2176 (1895).
- (185) TIEMANN: Ber. **30**, 598 (1897).
- (186) TIEMANN: Ber. **33**, 3703 (1900).
- (187) TIEMANN AND SCHMIDT: Ber. **31**, 883 (1898).
- (188) TIEMANN AND SEMMLER: Ber. **28**, 2141 (1895).
- (189) TOIVONEN, OSARA, AND OLLILA: Acta Chem. Fennica **6B**, 67 (1933).
- (190) TOLLENS: Ber. **14**, 1950 (1881).
- (191) VAN ARENDONK AND CUPERY: J. Am. Chem. Soc. **53**, 3184 (1931).
- (192) VAN ARENDONK, CUPERY, AND ADAMS: J. Am. Chem. Soc. **55**, 4228 (1933).
- (193) VAUBEL AND SCHEUER: Z. angew. Chem. **18**, 215 (1905).
- (194) VIGNON: Compt. rend. **110**, 534 (1890).
- (195) VINCENT AND DELACHANAL: Bull. soc. chim. [3] **3**, 681 (1890).
- (196) VORLÄNDER AND KOHLMANN: Ber. **32**, 1878 (1899).
- (197) WAGNER AND ERTSCHIKOWSKY: Ber. **29**, 881 (1896).
- (198) WALLACH: Ann. **275**, 149 (1893).
- (199) WALLACH: Ber. **30**, 427 (1897).
- (200) WALLACH: Ann. **345**, 142 (1906).
- (201) WALLACH: Ann. **379**, 187 (1911).
- (202) WALLACH: Ann. **381**, 83 (1911); **414**, 286 (1918).
- (203) WALLACH: Ann. **384**, 202 (1911).
- (204) WALLACH: Ann. **389**, 193 (1912).
- (205) WALLACH: Ann. **394**, 362 (1912).
- (206) WALLACH: Ann. **395**, 85 (1913); **397**, 194 (1913).
- (207) WALLACH: Ann. **408**, 190 (1915).
- (208) WALLACH AND BLEMBEL: Ann. **319**, 103 (1901).
- (209) WALLACH, COLLMANN, AND THEDE: Ann. **327**, 139 (1903).
- (210) WALLACH AND KEMPE: Ann. **329**, 99 (1903).

- (211) WALLACH AND KÖHLER: Ann. **339**, 110 (1905).
- (212) WALLACH AND KÖHLER: Ann. **339**, 113 (1905).
- (213) WALLACH AND RITTER: Ann. **381**, 92 (1911).
- (214) WARUNIS AND LEKOS: Ber. **43**, 654 (1910).
- (215) WEIDEL AND GRUBER: Ber. **10**, 1137 (1877).
- (216) WERNER: J. Chem. Soc. **85**, 1377 (1904).
- (217) WOLFF AND RÜDEL: Ann. **294**, 199 (1897).
- (218) WOODWARD AND FUSON: J. Am. Chem. Soc. **55**, 3472 (1933).
- (219) ZETZSCHE: Pharm. Zentralhalle **44**, 505 (1903).
- (220) ZINCKE: Ber. **26**, 322 (1893).
- (221) ZINCKE: Ber. **26**, 498 (1893).
- (222) ZINCKE: Ber. **26**, 501 (1893).
- (223) ZINCKE AND EGLY: Ann. **300**, 184 (1898).
- (224) ZINCKE AND FRANCKE: Ann. **293**, 147 (1896).
- (225) ZINCKE AND FUCHS: Ber. **25**, 2694 (1892).
- (226) ZINCKE AND FUCHS: Ber. **26**, 506 (1893).
- (227) ZINCKE AND FUCHS: Ber. **26**, 1666 (1893).
- (228) ZINCKE AND GERLAND: Ber. **21**, 2391 (1888).
- (229) ZINCKE AND GERLAND: Ber. **21**, 2396 (1888).
- (230) ZINCKE AND GÜNTHER: Ann. **272**, 243 (1893).
- (231) ZINCKE AND LÖHR: Ber. **25**, 2219 (1892).





# THE CHEMISTRY OF THE BILE ACIDS AND RELATED SUBSTANCES

HARRY SOBOTKA

*Department of Chemistry, The Laboratories of the Mount Sinai Hospital,  
New York City, New York*

*Received September 2, 1933*

*Revised to April 1, 1934*

## CONTENTS

I. Introduction.....	312
II. Occurrence of bile acids in nature.....	313
III. Methods of structural investigation.....	314
A. Oxidation; dehydrogenation.....	314
B. Hydrogenation; reduction.....	320
C. Dehydration.....	323
D. Thermic cyclization.....	324
E. Other reactions of keto groups.....	329
F. Reduction of carboxyl groups.....	331
IV. Systematic description of bile acids and their degradation products....	333
A. Hydroxycholic acids.....	333
B. Ketocholic acids.....	333
C. Acids with three rings.....	334
D. Acids with two rings.....	338
E. Acids with one ring.....	341
F. Steric isomerism.....	345
G. Sterols with longer side chains.....	346
H. Transitions. Modifications of the side chain.....	348
J. Degradation of side chain; hormones and carcinogenic products....	352
V. Chemical and physical properties of bile acids.....	354
A. Analytical methods; color reactions.....	354
B. Physical properties.....	355
C. Monomolecular layers.....	356
D. X-ray analysis.....	356
E. Dissociation constants.....	357
VI. Cholic acids.....	358
A. Coördination number.....	361
B. Coördination compounds and optical activity.....	362
C. Multiple coördination numbers.....	363
D. Tautomerism and coördination.....	363
E. Resorption in the intestinal tract.....	364
VII. Concentration and pH of bile.....	366
VIII. Physiological and pharmacological observations.....	366
IX. Conclusion.....	368

## I. INTRODUCTION

The knowledge of the constitution of the bile acids and the sterols has recently (1932) been revolutionized. The hypothetical structural formulas developed step by step through the continuous experimental efforts of Wieland, Windaus, and others during the last twenty years had to be fundamentally revised on the basis of new evidence. Adumbrated by certain findings of Diels, the new conception grew out of x-ray spectroscopy (13). Rosenheim and King "supplied the spark which has led to such a notable bonfire of unseemly molecular structures." (Robinson, in reference 51.) In the words of the English investigators, their contribution "would have been impossible but for the fundamental experimental work which forms the basis of present-day knowledge and for which we are indebted to the unremitting labors of Mauthner, Windaus, Diels, Wieland, Borsche, Schenck, and their collaborators" (156). It may be added that during the past two years the German and the British groups have assiduously translated most of the older evidence into terms consistent with the new conception. Thus, most discrepancies, unaccounted for by the old formulas, have vanished and a rich crop of additional facts harmonizes with the revised interpretation of the older.

The physiology of the bile acids and sterols has advanced steadily abreast with the progress of the medical sciences, especially experimental surgery and pharmacology, but one can hardly fail to observe the wide gap between these biological investigations and the chemical research. Yet the greater lucidity, recently attained in the chemistry of the bile acids, should dispel the reluctance of physiological chemists to link the remarkable chemical characteristics of these substances with their equally noteworthy place in the physiology and pathology of the animal organism. These bridges will be built along three lines of thought, two of which, while comparatively young, are sufficiently advanced to warrant their application. First, colloidal chemistry, or more properly, capillary chemistry, the importance of which for the extremely surface-active bile acids requires no comment. Second, the chemistry of molecular or coördination compounds, inaugurated by Werner, and applied to organic substances particularly by Pfeiffer. It was this concept which inspired Wieland to establish the "choleic acid principle" (200). The third idea which rather inheres in this special branch of organic chemistry, is that of the genetic relation between bile acids and sterols, presaged perhaps for the first time by Lachinov (111), and corroborated through a steadily converging series of investigations, culminating in the interconversion of bile acids and sterols, and stimulated by the discovery of such physiologically active substances of related structure as vitamin D, the sex hormones, and certain carcinogenic substances.

At this juncture, a review must reveal a vista from the point that has

been reached, giving proofs for the new ideas rather than explanations for past errors. This does not preclude the value and the necessity of consulting the older sources for the student who wishes to acquaint himself with the innumerable experimental facts on record, but is eager to interpret them in the light of modern knowledge.<sup>1</sup>

## II. OCCURRENCE OF BILE ACIDS IN NATURE

Bile acids are the preponderant constituent of bile, the secretion of the liver in vertebrate animals. They occur in conjugated form, i.e., in a peptide-like linkage with the amino acids glycine and taurine. Toad-poison, bufotoxin, contains arginine and suberic acid in conjugation with bufotalin which is a bile acid derivative. The lowest order of fish excretes in its bile the "biloid" scymnol as a sulfuric acid ester. The significance of conjugation, which imparts greater water-solubility, is beyond the scope of this review. It may be said, however, that the liver uses conjugation also for purposes of elimination, e.g., in the case of benzoic acid. Synthetic conjugation with glycine and taurine was verified for cholic acid by Bondi and Müller (18) using Curtius' method (cholic acid ethyl ester  $\rightarrow$  cholyl hydrazide  $\rightarrow$  cholyl azide  $\rightarrow$  cholyl glycine or cholyl taurine), and for deoxycholic acid by Wieland (202).

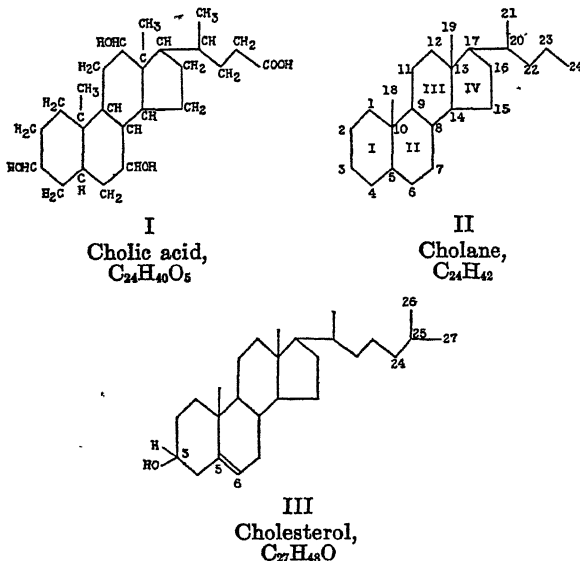
The ratio of taurine to glycine derivatives in bile varies from species to species; hog bile contains practically no taurine, while the bile of dogs, fish, snakes, etc., is devoid of glycine. Human bile and ox bile contain approximately equal amounts of tauro- and glyco-bile acids.

The cleavage of the peptide bond is presumably effected in nature by an enzyme (cf. 79, 131). In the laboratory, it is accomplished by alkaline hydrolysis. The resulting nitrogen-free unconjugated, or simply "free," bile acid consists of a mixture of acids in proportions which vary with species and with local and seasonal factors. At least half a dozen natural free bile acids have been isolated. The most frequent are cholic acid and deoxycholic acid; hyodeoxycholic acid prevails in hog bile, and chenodeoxycholic acid in goose bile.

### *Constitution*

Cholic acid (I) is a saturated trihydroxymonocarboxylic acid derived from the hydrocarbon cholane,  $C_{24}H_{42}$  (II), which is eight hydrogen atoms poorer than the corresponding paraffin. Therefore, it must have four rings. Its formula is as follows (155, 229, 267).

<sup>1</sup> A simplified nomenclature in this field was proffered in references 158 and 161, and has been followed in the present review. Summaries of recent chemical developments are given in references 161, 94, and 267; discussions of the biochemical aspects may be found in references 178 and 145.



The numbering of the carbon atoms and of the rings is conventional; the position of the two methyl groups is not definitely proven. Cholesterol (III), the chief representative of animal sterols, is a secondary mono-unsaturated tetracyclic alcohol,  $C_{27}H_{48}OH$ , whose formula will be discussed simultaneously with that of the bile acids (155; cf. 185).

### III. METHODS OF STRUCTURAL INVESTIGATION

Before surveying the arsenal of reactions which serve as tools in the structural investigations, one should mention that the development of organic microanalysis was a direct outcome of Pregl's researches on the bile acids (120). It was essential for the achievements of bile acid chemistry, where a long sequence of reactions, despite fair yields, reduces kilograms of starting material to mere grams of degradation products holding promise of unequivocal information. On the other hand, physical methods, such as measurements of unimolecular surface films (1, 2), x-ray spectroscopy (13), and ultra-violet absorption spectra (133, 157), have guided the structural research at important crossroads.

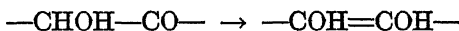
#### A. Oxidation; dehydrogenation

The foremost place among chemical reagents employed in disentangling the carbon skeleton of bile acids and sterols is occupied by oxidizing agents.

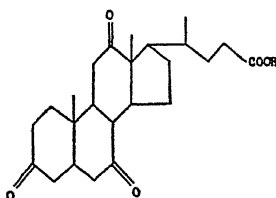
Chromic acid,  $CrO_3$ , in sulfuric or acetic acid solutions of high concentrations, is the most specific oxidant for the conversion of secondary alcoholic hydroxyl groups into keto groups, when used at low temperatures.

The classical example for its application to the present subject remains the dehydrogenation of cholic acid by Hammarsten (87) to a triketo acid known as dehydrocholic acid (IV). The reaction products of this method are regularly referred to as dehydro acids. Their functional nitrogenous derivatives, such as oximes, semicarbazones, and hydrazones, are extremely useful for the characterization of the parent substances. As will be seen later, these keto acids lend themselves better to reduction than the hydroxy acids and are indispensable intermediates in the replacement of secondary alcohol groups by methylene groups. Their isolation is desirable, although not imperative, when the destruction of a ring by stronger reagents, such as chromium trioxide at higher temperatures, nitric acid, or bromine, is contemplated. These three reagents break up the grouping  $-\text{CH}_2-\text{CO}-$  into  $-\text{CO}_2\text{H}|\text{HO}_2\text{C}-$ , sometimes via the diketone  $-\text{CO}-\text{CO}-$  which has been isolated in rare instances (e.g., the bilisoidanic acid of Schenck (167)).

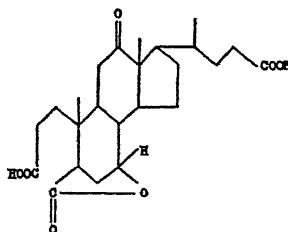
The mechanism of oxidation by bromine, which was first attempted by Landsteiner (117), has been studied extensively, and many mono- and dibromo derivatives of the  $-\text{CHBr}-\text{CO}-$  and  $-\text{CHBr}-\text{CO}-\text{CHBr}-$  type have been isolated, especially by Wieland and Dane (44, 226). These  $\alpha$ -bromoketo substances undergo numerous reactions, such as replacement of bromine by hydroxyl, and, depending on the presence of neighboring methylene groups, loss of hydrogen bromide yielding  $-\text{CH}=\text{CH}-\text{CO}-$ , and enolization yielding enediols.



The action of bromine on hydroxyl groups is peculiarly specific; in the case of cholic acid, bromine will disrupt ring I between carbon atoms  $C_3$  and  $C_4$  as outlined above, it will oxidize the hydroxyl group on  $C_{12}$  to a carbonyl, but it will leave the hydroxyl on  $C_7$  untouched so that the intervening lactonization of this hydroxyl with the carboxyl  $C_4$  leads mainly to the formation of biliobanic acid (V) (144, 212).

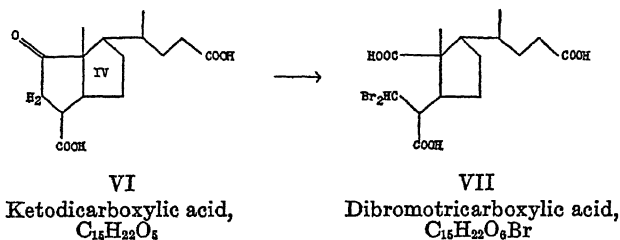


IV  
Dehydrocholic acid,  
 $\text{C}_{24}\text{H}_{44}\text{O}_5$

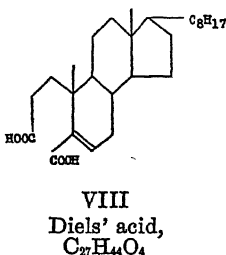


V  
Biliobanic acid,  
 $\text{C}_{24}\text{H}_{40}\text{O}_7$

The action of bromine on methyl ketones, resulting in the splitting-off of bromoform, finds its counterpart in the effect of bromine on a keto derivative (VI), which is converted into an  $\omega$ -dibromocarboxylic acid (VII) (224).

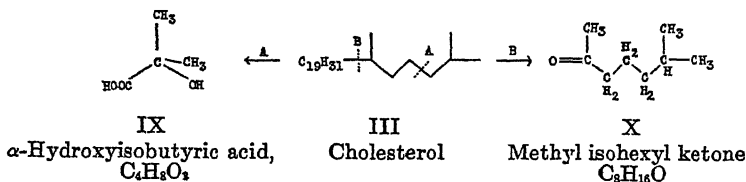


The use of potassium hypobromite on cholesterol leads to a clean scission of ring I between  $\text{C}_3$  and  $\text{C}_4$  (VIII) (46).

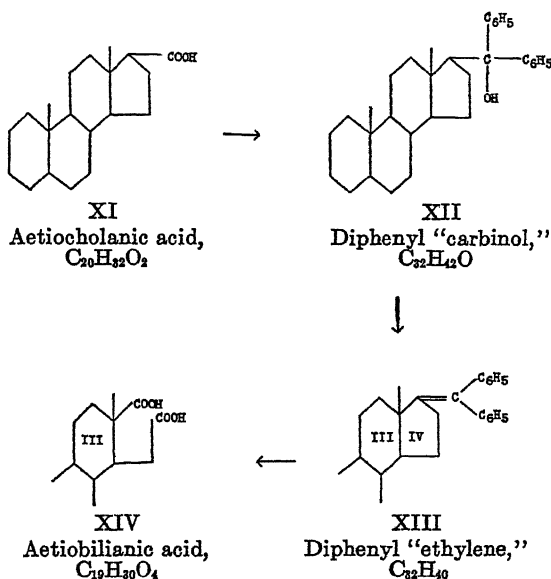


The most vigorous, but least lucid oxidations are effected with nitric acid of specific gravity 1.40–1.42 or fuming nitric acid of specific gravity 1.52. They lead to the opening of two or all of the three rings (I, II, and III) and, while the yield of a specific desired reaction product is not so advantageous as a rule, nitric acid still deserves its place in this field. In some instances, intermediates containing a nitro group have been isolated; for instance, nitrocholestenone, where the nitro group is attached to an "aromatic" carbon atom.

Nitric acid, and subsequently chromium trioxide, were applied by Windaus (246, 247, 249) in the destructive oxidation of the side chain of cholesterol, leading to the isolation of acetone,  $\alpha$ -hydroxyisobutyric acid (IX), and methyl isohexyl ketone (X) (248).

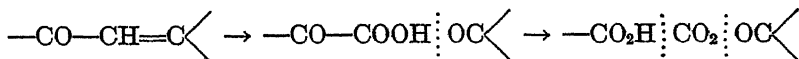


Chromium trioxide was also the oxidizing agent used by Wieland, Schlichting, and Jacobi (221) in the stepwise Grignard degradation of the side chain in cholanic acid, and it was with this reagent that the first and so far only inroad was made into ring IV after the side chain was replaced in the last step by  $\text{>C}=\text{C}(\text{C}_6\text{H}_5)_2$  (XI to XIII), leading to the formation of two new carboxyl groups in aetiobilianic acid (XIV). For the formation of *n*-butane-1,3,3'-tricarboxylic acid from Wieland's diketodicarboxylic acid (CXXXII from CII), see p. 344.



Potassium permanganate differs from the previous oxidizing agents in its action on keto groups. It will sever a carbonyl group from a tertiary rather than from a secondary carbon neighbor. In the model substance  $\alpha$ -methylcyclopentanone, it produces acetylbutyric acid ( $\delta$ -ketocaproic acid) rather than methylpimelic acid (196). This type of ring cleavage, while especially frequent with permanganate (Wieland's diketodicarboxylic acid (CII),  $\text{C}_{22}\text{H}_{34}\text{O}_6$ , from pyrodeoxybilianic acid (CI)), has been encountered occasionally in oxidations with nitric acid.

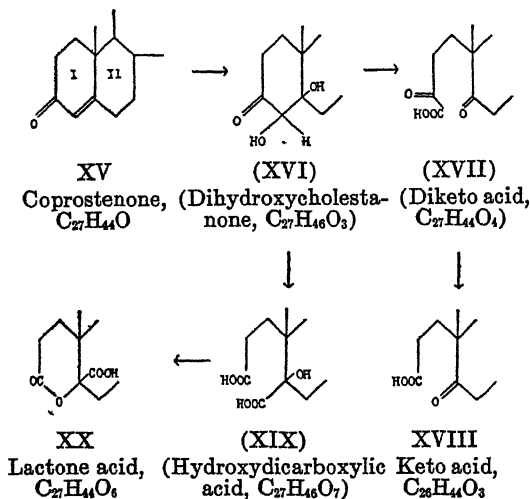
Neutral permanganate, for instance, in acetone solution also leads to the formation of a keto acid in cases of  $\alpha,\beta$ -unsaturated ketones like coprostenone



through a series of hypothetical intermediates (XV to XVIII), which at

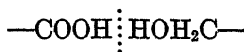


the same time may serve to explain the formation of the lactone acid,  $C_{27}H_{44}O_6$ , in the example given below (XIX to XX) (185, 245).

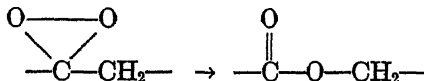


Cyclic  $\alpha$ -hydroxy acids, on treatment with permanganate in sulfuric acid, will yield cyclic ketones through the loss of the elements of formic acid (197); see p. 330, formula for cilianic acid (LXIII).

While potassium permanganate is the traditional oxidant for aliphatic double bonds (Baeyer's reagent), the cleanest method for the introduction of two vicinal hydroxyl groups uses hydrogen peroxide, for instance, in the preparation of cholestantriol (143, 253); benzoyl peroxide may also be used (198). A peculiar reaction is the oxidation of cyclic ketones,  $-\text{CO}-\text{CH}_2-$ , by ammonium persulfate (9), yielding a primary hydroxy acid



presumably (244) via

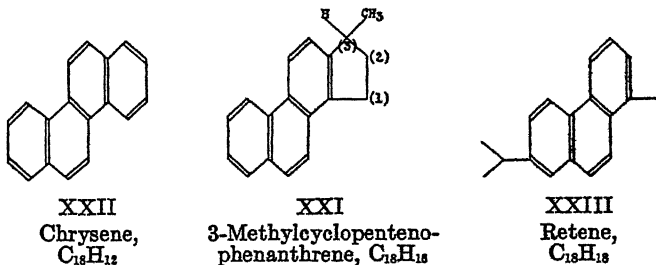


The appearance of acetone in peroxide oxidations is not confined to isopropyl derivatives; acids with secondary methyl groups, such as  $\alpha$ -methylglutaric acid, also form acetone (224).

Ozone has been applied to the double bond of cholesterol by Dorée and Gardner (54, 56) and Dorée and Orange (57; cf. 90). Butenandt used ozone in his studies on pregnandiol (30). Levorotatory isopropylmethyl-

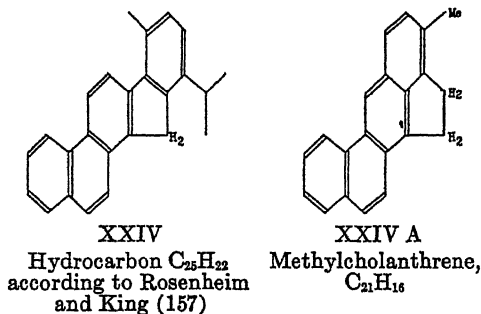
acetaldehyde has been obtained from ergosterol, and isopropylethylacetaldehyde from stigmasterol by the action of ozone upon the double bond between  $C_{22}$  and  $C_{23}$  in these sterols (150, 82, 83).

Dehydrogenation proper, resulting in the formation of double bonds, has been studied especially by Diels (47). When dehydrogenation of saturated hydrocarbons and their halogen derivatives with palladized charcoal catalysts proved unsatisfactory, Diels resorted to sulfur (cf. 190). As this element tends to insert itself into cyclic structures at the elevated temperatures of the reaction, he replaced it by selenium. Dehydrogenation of cholesterol by selenium yielded two hydrocarbons, 3-methylcyclopentenophenanthrene (XXI),  $C_{18}H_{16}$ , and  $C_{25}H_{22}$  (or  $H_{24}$ ), while palladized charcoal on cholesterol or selenium on cholic acid led to a hydrocarbon,  $C_{18}H_{12}$ , identified by Diels (48, 49, 50) as chrysene (XXII). These aromatic and semiaromatic hydrocarbons contain the tetracyclic skeleton of cholesterol. The methyl groups attached to tertiary carbon atoms were eliminated as methane, or shifted from  $C_{13}$  to  $C_{17}$ , in the course of dehydrogenation (aromatization), and the side chain was either dropped, giving rise to paraffins such as propane, hexane, and isoöctane, or a new ring was formed from it. These aromatizations of polycyclic compounds are not always complete, depending on the catalyst employed. It has been pointed out by Rosenheim and King (156), "that dehydrogenation with selenium leads to the fully aromatic hydrocarbon,  $C_{18}H_{12}$ , i.e., chrysene, when applied to cholic acid with its three alcoholic hydroxyl groups, symmetrically disposed in three rings, whilst the same process stops at the partially saturated hydrocarbon,  $C_{18}H_{16}$ , in the case of the sterols (cholesterol and ergosterol), where the point of attack is limited to a single hydroxyl group in one of the outside rings." Additional evidence along these lines may be expected from the studies on abietic acid,  $\beta$ -pimaric acid, the hydrocarbon retene (XXIII), and other partly hydrogenated phenanthrene derivatives (91).



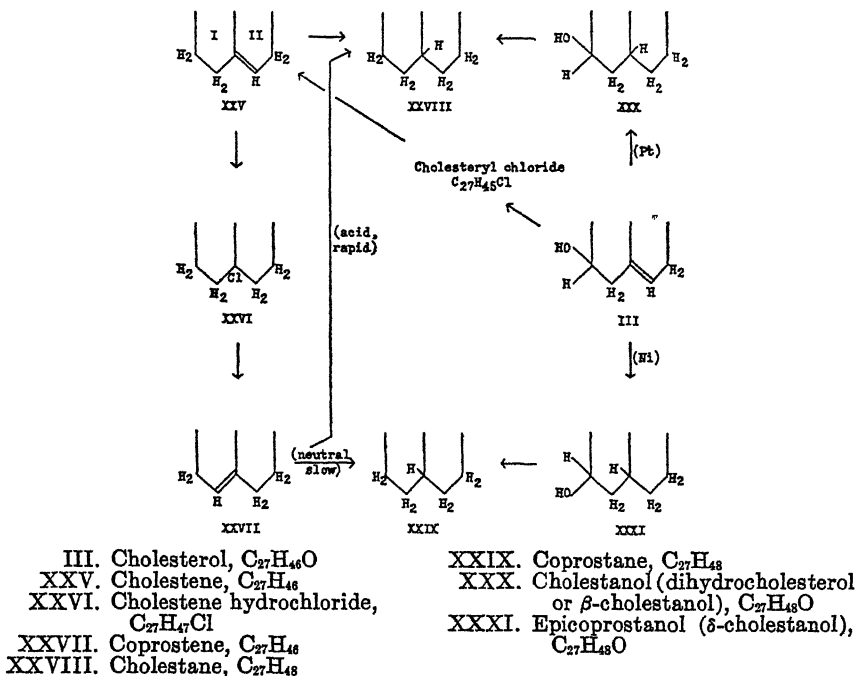
The belief that the aromatic products obtained by these methods are not due to a complicated pyrosynthesis, but represent the true carbon skeleton of cholesterol and of the bile acids, has been corroborated (e.g., see reference

147). Ruzicka, Goldberg, and Thomann (164) have cast some doubt on the identity of  $C_{18}H_{12}$  from cholesterol with chrysene, but spectroscopic evidence (50) supports it. The identity of  $C_{18}H_{16}$  with 3-methylcyclopentenophenanthrene has been confirmed by comparison of the substance with synthetic cyclopentenophenanthrene (40), 1-methyl- and 2-methylcyclopentenophenanthrene (108, 164), and 3-methylcyclopentenophenanthrene (10, 89, 14). The oxidation of  $C_{25}H_{22}$  with chromium trioxide yields the monoketone  $C_{25}H_{20}O$  (XXIV). Aromatic six-membered rings give rise to quinones under these conditions, whilst five-membered rings of the fluorene type form monoketones. These facts, suggesting the original presence of a five-membered ring, are in agreement with other experimental evidence concerning ring IV of sterols and bile acids (157). Another pentacyclic derivative, dehydronorcholene (LX), whose formation from bile acids will be derived on p. 329, yields on dehydrogenation a hydrocarbon,  $C_{21}H_{16}$ , methylcholanthrene (236, 39), which was found by the British workers to be a most active carcinogenic substance (cf. p. 354 and formula XXIV A).



### B. Hydrogenation; reduction

Among the various methods of hydrogenation, one has to enumerate catalytic reduction by hydrogen in the presence of palladium, platinum, or nickel at normal or elevated temperature and pressure. These catalysts have been used widely, especially in the hydrogenation of double bonds. Several members of the sterol series, such as cholesterol itself, are endowed with double bonds. If one of the double bonded carbon atoms is tertiary, i.e., if it carries no hydrogen atom, then the addition of  $H-H$  will give rise to a new center of asymmetry and two dihydro compounds may theoretically be formed (cf. 173). In reality, this will occur with some hydrogenating agents, whilst others will lead to the preferential formation of one isomer. The interrelations between cholestene (XXV), coprostene (pseudocholestene) (XXVII), cholestane (XXVIII), and coprostane (pseudocholestane) (XXIX) may serve as an example.



The scheme illustrates at the same time the shift of double bonds through intermediary addition and subtraction of HCl, first observed by Mauthner (130). Another example for this reaction is given by Rosenheim and King (156), who, in unpublished experiments with Starling, converted apocholic acid by hydrochloric acid into the structural isomer dihydroxycholenic acid.

The hydrogenation of cholesterol itself in the presence of platinum yields cholestanol (dihydrocholesterol), according to Willstätter and Meyer (243), while the reduction with nickel as catalyst results in a mixture ( $\gamma$ -cholestanol) of cholestanol ( $\beta$ -cholestanol) and epicoprostanol (" $\delta$ -cholestanol").

Other unsaturated compounds are formed especially from bile acids by the loss of H—OH from the grouping  $R_1R_2CH-COHR_2R_3$ , resulting in  $R_1R_2C=CR_3R_4$ . If neither  $R_1$  nor  $R_2$  represents a hydrogen atom, then atom  $C$  must have been a center of asymmetry in the hydroxy compound. This asymmetry disappears on dehydration and is restored through the subsequent hydrogenation of the ethenoid linkage. Again, hydrogenation potentially yields two stereomers. In the case of multiple unsaturation, as in ergosterol among natural compounds or in cholatrienic acid among synthetic derivatives, at least eight stereomers may form according to the

theory. It seems, however, that the hydrogenation of such entwined polycyclic structures is governed by certain laws of mutual induction and leads to the preferred formation of a very limited number of products, frequently just of a single isomer.

The course of the hydrogenation of keto groups, particularly in the aforementioned dehydro acids, depends likewise on the method used. Electrolytic reduction of dehydrocholic acid, or action of sodium amalgam or aluminum amalgam in ether-benzene, leads to 3-hydroxy-7,12-diketo-cholanic acid (reductodehydrocholic acid), while Willstätter's platinum catalyst yields 3,7-dihydroxy-12-ketocholanic acid plus 3,7,12-trihydroxycholanic acid (cholic acid), which is the sole reaction product when platinum at 80–90°C. and hydrogen at 3 atmospheres pressure is used (Skita's method). One of the most successful hydrogenation methods is that developed by Clemmensen in the laboratories of Parke, Davis and Co. (35). Clemmensen used zinc amalgam in the presence of hydrogen chloride in glacial acetic acid for the reduction of keto groups in aliphatic, hydroaromatic, aromatic, and polycyclic compounds. For example, zinc amalgam will reduce dehydrocholic acid to cholanic acid. While, according to Clemmensen's experience, aromatic hydroxyl groups are left intact (aromatic oxyketones  $\rightarrow$  phenols), hydroaromatic secondary alcohol groups are reduced; for instance, in the case of reductodehydrocholic acid the hydroxyl group, as well as both keto groups, is reduced, yielding cholanic acid.

The complete reduction of carbonyl, with secondary alcohol groups remaining unchanged, can be accomplished by the procedure of Kishner (106) and of Wolff (268). The hydrazone or preferably the semicarbazone is prepared; absolutely anhydrous sodium ethylate will then replace the  $=N-NH_2$  or  $=N-NH-CO-NH_2$  by  $H_2$ , without any reaction on the hydroxyl group. Double bonds are left intact both by Kishner-Wolff's and by Clemmensen's method. The degree of resistance of keto groups against hydrogenation varies with the position of the carbonyl. Those in ring I, especially on  $C_3$ , are the most reactive; those in ring III, on  $C_{12}$ , the least reactive. This factor can be utilized in various ways for a great variety of preparative purposes in combination with partial dehydrogenation and dehydration. Parallel differences of reactivity will also be found in the tendency of the various carbonyl groups to form oximes and other functional derivatives.

Sodium metal in alcoholic solutions and sodium ethylate reduce carbonyl groups to carbinol. Their occasional hydrogenating action on double bonds in unsaturated alcohols like cholesterol seems to be due to intervening rearrangements into the respective saturated ketone (e.g., cholesterol  $\rightarrow$  cholestanone  $\rightarrow$   $\gamma$ -cholestanol).

The replacement of halogen by hydrogen offers no peculiarities in the case of halogenated bile acid derivatives.

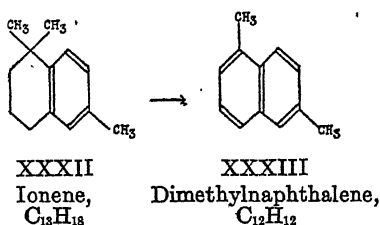
### C. Dehydration

A considerable number of unsaturated acids have been prepared from the various natural hydroxy acids of bile by distillation in a high vacuum. The removal of the constituents of water from  $\text{—CHH—CHOH—}$  may take place in several stages, depending on the number of secondary alcohol groups in the molecule. Thus, lithocholic acid (3-hydroxycholanolic acid) gives  $\Delta^{2,3}$ - or  $\Delta^{3,4}$ -cholenic acid, the several dihydroxycholanolic acids, (3,7- or chenodeoxycholic acid, 3,6- or hyodeoxycholic acid, 3,12- or common deoxycholic acid, and the synthetic 7,12- or isodeoxycholic acid) yield a series of respective choladienic acids, while cholic acid is dehydrated to cholatrienic acid (m. p. 163–164°C.), accompanied by several isomers in minor amounts (173). All these acids can be hydrogenated catalytically to the mother substance cholanolic acid, which in fact was obtained for the first time via cholatrienic acid (199, 201).

Milder dehydrating agents like zinc chloride, concentrated hydrochloric acid, or potassium bisulfate lead by partial dehydration to unsaturated hydroxy acids such as dihydroxycholenic acids or hydroxycholadienic acids. The most important representatives of this type of compounds are 3,12-dihydroxycholenic acid and apocholic acid. Both were obtained simultaneously by Boedecker, and the former was transformed by catalytic hydrogenation into deoxycholic acid according to Boedecker's intention (17). This was the first transition from cholic acid to its most important companion acid. Claims of biochemical reduction of cholic acid to deoxycholic acid by microorganisms have never been experimentally substantiated. The difficulty of explaining the isomerism between dihydroxycholenic acid and apocholic acid led into various blind alleys until it was solved through the establishment of the new formula. Apocholic acid reacts with bromine and other oxidizing agents, but, although the presence of a double bond is betrayed by a yellow coloration with tetranitromethane, it displays absolute resistance against hydrogenation. This behavior finds a most plausible explanation (233) in the assumption that the double bond has migrated from  $\Delta^{7,8}$  to  $\Delta^{8,9}$  or  $\Delta^{8,14}$ . This view is supported by the analogy of the formation of  $\Delta^{9,10}$ -octalene (octahydronaphthalene) from  $\alpha$ -decalol (97). This octalene likewise escapes hydrogenation and the same holds for hexadecahydrochrysene (28). This gradation in affinity towards hydrogen finds its counterpart in the stepwise dehydrogenation with selenium discussed previously (156; cf. 105).

Ergosterol through a peculiar dimeric intermediary compound may be converted into neorgosterol,  $\text{C}_{27}\text{H}_{39}\text{OH}$  (99). While its formula calls for

four double bonds, only one can be detected by oxidation with benzoyl peroxide or catalytic hydrogenation; the other three double bonds have entered into a stable aromatic configuration presumably in ring II, as supported by the identification of 1,2,3,4-benzenetetracarboxylic acid, a remnant of its destructive oxidation with nitric acid. Toluenetetracarboxylic acid has been obtained by nitric acid oxidation of ergosterol itself. The methane originating by the cleavage of  $\text{CH}_3$  and H from vicinal carbon atoms has been recovered by Inhoffen. This aromatization through loss of a methyl group follows the model of ionene (XXXII), which is converted by heating with sulfur into 3,8-dimethylnaphthalene (XXXIII) (163).



Tendency to aromatization through loss of  $\text{CH}_3\text{H}$  is characteristic for apocholic acid and its derivatives (233). Considerable amounts of methane are recovered when such substances are heated to  $320\text{--}340^\circ\text{C}$ . in an inert atmosphere such as carbon dioxide.

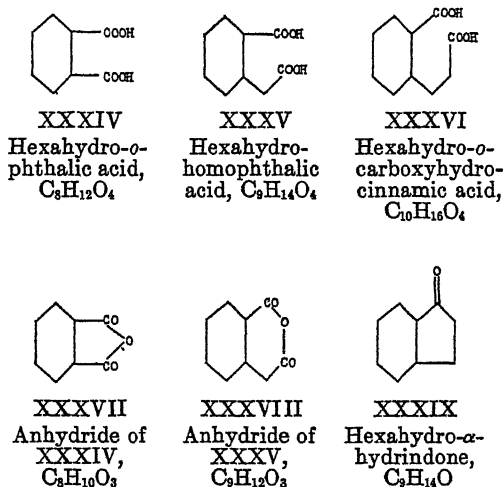
Cruder methods of dehydration lead to the formation of anhydrides and to polymerization of bile acids to amorphous insoluble masses long known as "dyslysins." The dehydration of keto acids gives rise to condensations between the carbonyl group of one molecule with a reactive methylene group of another molecule, but the substances mentioned in this paragraph are without significance for structural questions.

#### *D. Thermic cyclization*

Dehydration by milder methods of keto and other polycarboxylic acids derived from bile acids, takes an important place in constitutional analysis of the carbon skeleton and the position of hydroxyl groups. The interpretation of these reactions is based on two rules borrowed from simpler model reactions, the so-called Blanc's rule and Brecht's rule.

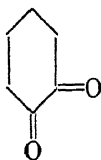
Blanc (15) observed that glutaric acid on treatment with acetic anhydride forms an intramolecular anhydride, while adipic and pimelic acids subjected to the same treatment are converted into cyclopentanone and cyclohexanone, respectively (cf. 8). This is only an example of the well-known tendency of five- and six-membered rings to form whenever cy-

clization occurs. Blanc's experience was established as a "rule" by the German investigators, and they noticed that similar cyclizations can be frequently accomplished without the use of acetic anhydride by mere application of heat in a high vacuum. This rule can be applied to a great many substitution products of aliphatic dicarboxylic acids, e.g., in the terpene series to the formation of camphononic acid from homocamphoronic acid (118). Windaus, Hückel, and Revere (256) studied the cyclization of hydroaromatic dicarboxylic acids, such as hexahydro-*o*-phthalic acid (XXXIV), hexahydrohomophthalic acid (XXXV), and hexahydro-*o*-carboxyhydrocinnamic acid (XXXVI) in the presence of acetic anhydride. The two former behave like succinic and glutaric acid, yielding anhydrides (XXXVII, XXXVIII) with five- and six-membered rings, while the cyclohexane-1-carboxylic acid-2-propionic acid (XXXVI) is transformed into hexahydro- $\alpha$ -hydrindone (XXXIX).

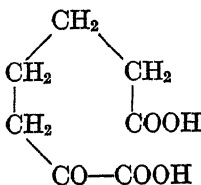


Adickes in Wieland's laboratory studied thermic reactions of  $\alpha$ -ketopimelic and  $\alpha$ -ketosuberic acids (4). The tendency towards the formation of cyclic diketones by loss of carbon dioxide and water by heating in nitrogen was greater in the case of ketopimelic acid (XL)  $\rightarrow$  cyclohexanedione (XLI) than for ketosuberic acid (XLIII)  $\rightarrow$  cycloheptanedione (XLIV). On the other hand, the reaction involving the loss of carbon dioxide, carbon monoxide, and water in the presence of concentrated sulfuric acid is realized to a higher percentage of the theoretical value in ketosuberic acid yielding cyclohexanone (XLV) than with the lower homolog (XLII). Thus in any event, reactions leading to the formation of six-membered rings are most prone to occur.





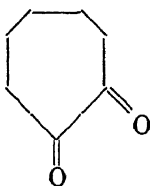
XLI  
Cyclohexanedione,  
 $C_6H_8O_2$



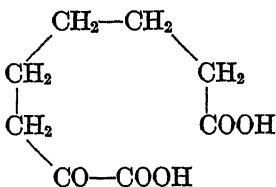
XL  
Ketopimelic acid,  
 $C_7H_{10}O_5$



XLII  
Cyclopentanone,  
 $C_5H_8O$



XLIV  
Cycloheptanedione,  
 $C_7H_{10}O_2$



XLIII  
Ketosuberic acid,  
 $C_8H_{12}O_5$

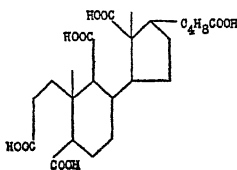


XLV  
Cyclohexanone,  
 $C_6H_{10}O$

From these observations, the following general rule was formulated: Formation of anhydrides that can be "saponified" to the original acid with greater or lesser ease speaks for a 1,4- or 1,5-position of the participating groups. Cyclic ketones originate from 1,6-, 1,7- and possibly 1,8-dicarboxylic acids by loss of carbon dioxide and water. That one of two carboxyl groups in the 1,3-position will be expelled at moderately elevated temperature, is to be expected from substituted malonic acids.

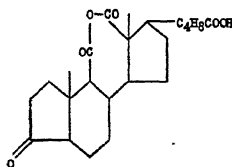
As such ketonic ring structures may be opened by oxidizing agents as outlined in pp. 315 to 318, oxidation alternating with thermic cyclization allows stepwise degradation as a general method of structural investigation. The great usefulness of this method and the agreement of its results with conceptions derived from other evidence led to a general acceptance of the above interpretation. This entailed severe misconceptions in those cases where the diagnostic value of the method was founded on analogy only. These errors were recognized by Rosenheim and King, who pointed out that the formation of anhydrides by thermic cyclization does not in every case exclude 1,6-position of the carboxyl groups involved. The anhydride formation, in substances like choloidanic acid (XLVI  $\rightarrow$  XLVII) or the isomer of lithobilianic acid, with cleavage of ring III (XCII  $\rightarrow$  XLVIII) led Windaus, Wieland, and others to the assumption that anhy-

dride formation is adequate proof for the 1,5-position of the two carboxyl groups involved. Thus the ring, now designated as ring III, was assumed to be a five-membered ring which, because of the obvious presence of  $-\text{CHOH}-\text{CH}_2-$  or  $-\text{CO}-\text{CH}_2-$ , could contain but three tertiary carbon atoms. Thilobilianic acid likewise forms an anhydride instead of a ketone, although ring II is six-membered ( $\text{XCI} \rightarrow \text{XLIX}$ ).



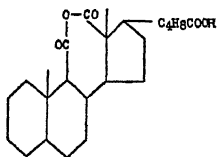
XLVI

Choloidanic acid,  
 $\text{C}_{24}\text{H}_{36}\text{O}_{10}$



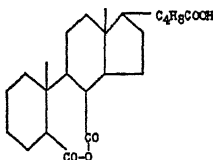
XLVII

Anhydride of pyrocholoidanic  
acid,  $\text{C}_{22}\text{H}_{32}\text{O}_6$



XLVIII

Anhydride of acid XCII,  
 $\text{C}_{24}\text{H}_{36}\text{O}_8$

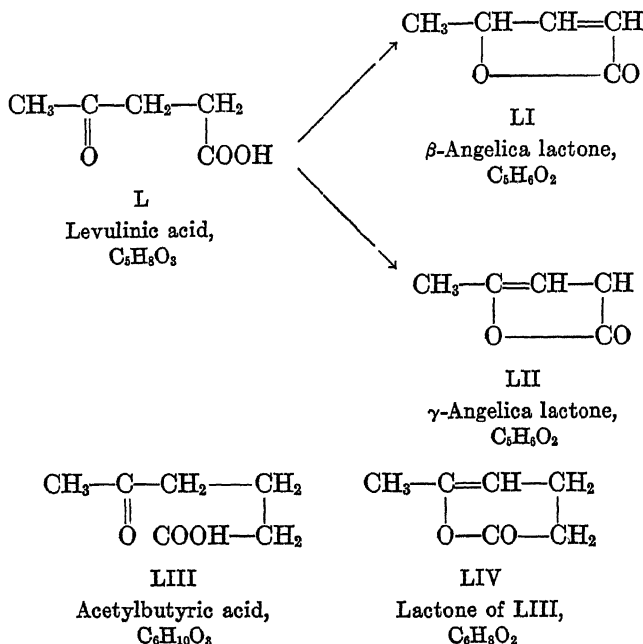


XLIX

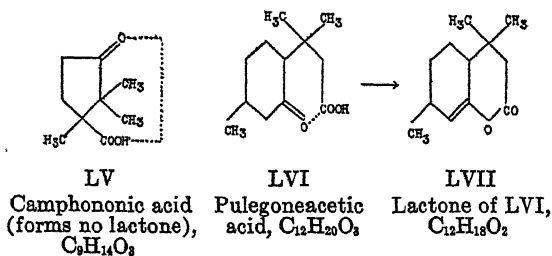
Anhydride of thilobilianic  
acid (XCI),  $\text{C}_{24}\text{H}_{36}\text{O}_8$

This in turn distracted attention from phenanthrene or chrysene systems as structural possibilities. An analogy in the aromatic series was unfortunately overlooked: diphenic acid forms an anhydride at  $270^\circ\text{C}$ . and no trace of the yellow fluorenone can be discerned in the distillate. This ketone forms, however, above  $320^\circ\text{C}$ . Hydroaromatic derivatives of diphenic acid have not yet been investigated, but it seems probable that the formation of seven-membered anhydride rings involves less distortion than that of five-membered ketone cycles.

The other regularity frequently encountered in this field pertains to the thermic cyclization of  $\gamma$ - and  $\delta$ -keto acids, giving rise to the formation of unsaturated lactones of the angelica lactone type. Windaus and Bohne (262) devoted a very comprehensive study to such unsaturated lactones of aliphatic and terpenoid character and deduced the following rules: (a) The carbonyl group must be in the  $\gamma$ - or  $\delta$ -position relative to the carboxyl group (L to LIV).



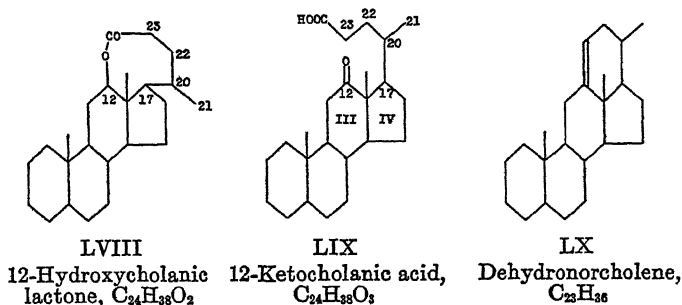
(b) There must be a hydrogen atom attached to the  $\alpha$ -carbon, i.e., the one next to the carboxyl group. (c) The keto group, when a member of a ring, must be in the  $\alpha$ -position to that carbon atom of the ring from which the carboxyl-containing side chain starts. Otherwise, the present ring and the newly formed lactone ring would share three carbon atoms. As the double bond must be situated either in the  $\alpha$ ,  $\beta$ - or in the  $\beta$ ,  $\gamma$ -position relative to the carbonyl group, this would lead to a conflict with the so-called Bredt rule. Bredt (29) observed that in terpenes with a 1,3-bridge, i.e., with two rings sharing three carbon atoms, no double bond could issue from either of the two pivotal atoms. Such acids, e.g., camphononic acid (LV), volatilize without decomposition, while the related pulegone-acetic acid is easily converted into an unsaturated lactone (LVI  $\rightarrow$  LVII) (193).



(d) The participation of the carboxylic side chain in a second cyclic structure also may cause steric hindrance against lactonization. (e) In polycarboxylic acids, reactions discussed previously as customary between pairs of carboxyl groups may successfully compete with lactonization. Thus, hydoxybiliaric acid forms an unsaturated lactone at 240–275°C., while temperatures above 300°C. foster the synthesis of a saturated pyroketone.

The newly formed ring in such unsaturated lactones, as well as in the various pyro derivatives of the anhydride and the cyclanone type, usually corresponds to one of the original rings present in the hydroxy and dehydro derivative. But there are instances where one of the two ring fragments will undergo cyclization with the fragment of another ring. Examples of such "synthetic rings" may be found in the case of pyrobiloidanic acid, (CXXIX, CXXX on p. 343) and pyrocholoidanic acid (CVIII and CIX on p. 339), which form unsaturated lactones. The cyclopentanone ring in the acid  $C_{15}H_{22}O_5$  (VI), pp. 316 and 345) is formed from fragments of rings II and III which were left over in acid  $C_{16}H_{24}O_8$  (CXXXI, p. 344).

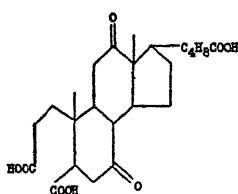
Other synthetic rings were observed in the formation of saturated lactones in the biliobanic series (see p. 337). Even the carboxyl group of the side chain may become involved, as exemplified in the lactonization of 12-hydroxycholanolic acid (LVIII) (219). Similar lactonizations seem to occur to a limited extent with polyhydroxycholanolic acids as deoxycholic acid. Intramolecular condensation between carbonyl in  $C_{12}$  and a methylene group of the side chain leads also to the formation of a fifth ring spanning from  $C_{17}$  in IV to  $C_{12}$  in III in the case of dehydronorcholene (LX) from 12-ketocholanolic acid (XLIX).



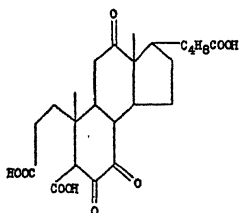
#### *E. Other reactions of keto groups*

The oxidation of a methylene group attached to a carbonyl group by permanganate results in the formation of a 1,2-diketone. These diketone substances easily undergo a rearrangement of the benzilic acid type in the presence of sulfuric acid, yielding a cyclic  $\alpha$ -hydroxy acid, while the number

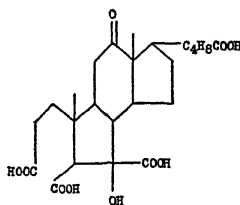
of carbon atoms within the ring is reduced by one. Thus, bilianic acid (LXI) is converted (LXII) into cilianic acid (LXIII) (22, 23).



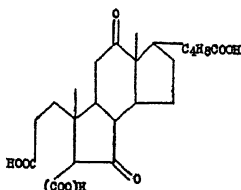
LXI  
Bilianic acid,  
 $C_{24}H_{34}O_8$



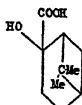
LXII  
Triketo acid,  
 $C_{24}H_{32}O_9$



LXIII  
Cilianic acid,  
 $C_{24}H_{34}O_{10}$



LXIV  
Ciloxanic acid,  
 $C_{22}H_{32}O_8$



LXV  
Nopinic acid,  
 $C_{11}H_{16}O_3$



LXVI  
Nopinone,  
 $C_{10}H_{14}O$

For one of the reactions of cilianic acid, we may again choose a model reaction from Wallach's extensive researches in the terpene group. Nopinic acid (LXV) loses carbon monoxide and water when treated with permanganate in sulfuric acid and a cyclic ketone, nopinone (LXVI), is obtained (197, see also p. 318). Analogously, cilianic acid (LXIII) may be converted into ciloxanic acid (LXIV) via an unstable  $\beta$ -keto acid which is decarboxylated in the course of its formation (24). The cyclopentanone ring of this acid is identical with that obtained by thermic cyclization of 6,7-dicarboxyl derivatives.

Other experimental attempts to investigate the nature of rings containing carbonyl followed Wallach's (195) application of the Beckmann oxime rearrangement to cyclic ketoximes (LXVII to LXIX).



LXVII  
Cyclopentanone oxime,  
 $C_5H_9ON$

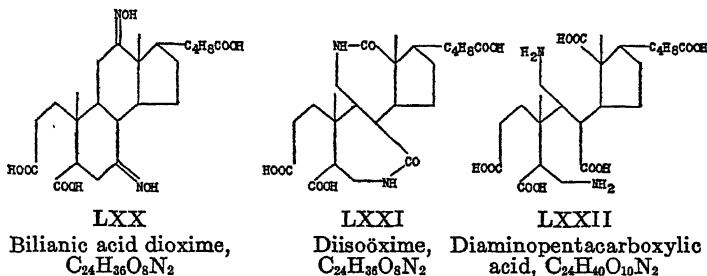


LXVIII  
Cyclopentanone isooxime,  
 $C_5H_9ON$   
(lactam of LXIX)



LXIX  
 $\delta$ -Aminovaleric acid,  
 $C_5H_{11}O_2N$

Hydroxylamine derivatives of dehydro bile acids were investigated in numerous cases by Schenck (166). The following compounds written in the modern formulation exemplify this rearrangement in the case of bilianic acid dioxime (LXX)  $\rightarrow$  diisoöxime (LXXI), which can be considered as a double lactam of a diaminopentacarboxylic acid (LXXII). The chemistry of these nitrogenous derivatives corroborates the findings by the methods previously discussed.

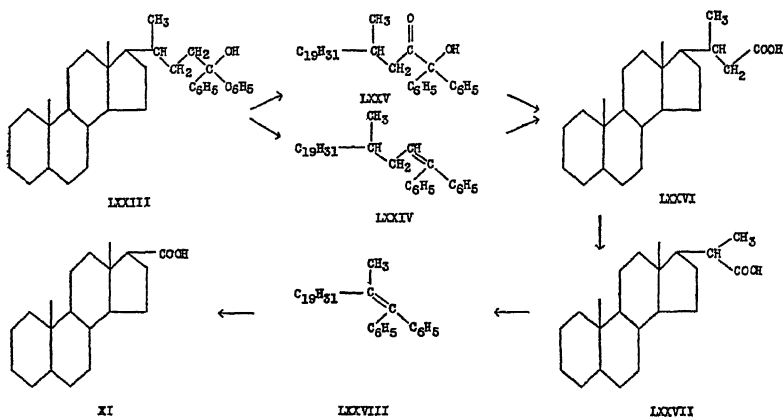


#### *E. Reduction of carboxyl groups*

Reduction of carboxyl groups formed by oxidative cleavage was investigated by Wieland (220) in the model compound hexahydrophthalic acid. Reduction with sodium and amyl alcohol according to Bouveault yielded the glycol and its furanoid oxide. Neither this reaction nor the conversion of hexahydrophthalic acid dimethyl ester by methylmagnesium iodide into the double dimethylcarbinol promised new insight when applied to bile acid derivatives.

A succession of Grignard reactions proved, however, a very valuable tool in the gradual destruction of the side chain, which was studied in great detail for the case of cholic acid by Wieland and his coworkers (221). The ethyl ester of this acid is converted by the appropriate Grignard reagent to the diphenyl-, dimethyl-, or diisopropyl- (cf. p. 348) carbinol. Such carbinols (LXXIII) yield on dehydration an unsaturated hydrocarbon (LXXIV) and on oxidation a ketocarbinol (LXXV). Both these derivatives, as well as the carbinol itself, can be oxidized with chromium trioxide to norcholic acid (LXXVI). This lower homolog is converted through an analogous set of reactions into the next homolog bisnorcholic acid (LXXVII). A third application of these methods leads to the unsaturated hydrocarbon diphenylbisanorcholene (LXXVIII). The oxidation of this substance brings about the simultaneous loss of two more carbon atoms. The resultant substance is called aetiocholic acid (XI). A fourth Grignard reaction applied to this substance leads to diphenyl-aetiocholene and, by oxidation, to a dicarboxylic acid, indicating that the nineteenth carbon atom is a member of a ring. This acid is aetiobilianic

acid, and the fact that it forms an anhydride by high vacuum distillation would indicate, according to "Blanc's rule," that this ring is a five-membered one.



LXXIII. Norcholanyldiphenylcarbinol,  $\text{C}_{36}\text{H}_{50}\text{O}$

LXXIV. Diphenylbischolene,  $\text{C}_{38}\text{H}_{48}$

LXXV. Ketocarbinol,  $\text{C}_{38}\text{H}_{48}\text{O}_2$

LXXVI. Norcholanic acid,  $\text{C}_{22}\text{H}_{38}\text{O}_2$

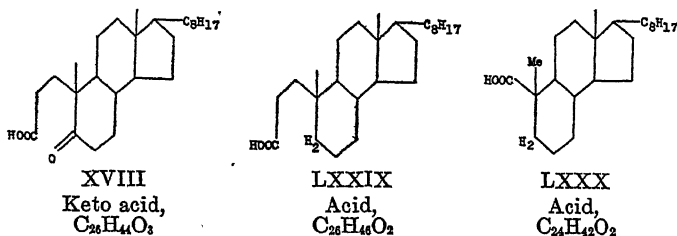
LXXVII. Bisnorcholanic acid,  $\text{C}_{22}\text{H}_{38}\text{O}_2$

LXXVIII. Diphenylbischolene,  $\text{C}_{32}\text{H}_{44}$

XI. Aetiocholanic acid,  $\text{C}_{20}\text{H}_{32}\text{O}_2$

This ring IV, which carries the side chain, is the one most resistant to any method of attack, since it never carries a hydroxyl group. Thus, oxidative methods will open up one, two, or all three of the other rings that carry hydroxyl groups in the naturally occurring bile acids, but ring IV will remain intact.

A similar series of Grignard reactions was applied by Tschesche (185) in the degradation of the keto acid (XVIII),  $\text{C}_{26}\text{H}_{44}\text{O}_3$ , obtained from coprostenone (XV). After reduction of the keto group to methylene (LXXIX), carbon atoms  $\text{C}_3$  and  $\text{C}_2$  were successively removed. The formation and the properties of the resulting acid (LXXX),  $\text{C}_{24}\text{H}_{42}\text{O}_2$ , corroborate the formula of Rosenheim and King. The great difficulty with which this acid is esterified supports the assumption that carbon  $\text{C}_{10}$  carries a methyl group ( $\text{C}_{18}$ ) besides the carboxyl group of  $\text{C}_1$ .



Grignard degradation of the side chain in a monocyclic (ring IV) acid will be further discussed below (p. 344).

#### IV. SYSTEMATIC DESCRIPTION OF BILE ACIDS AND THEIR DEGRADATION PRODUCTS

##### *A. Hydroxycholanolic acids*

At least seven hydroxylated cholanolic acids have been isolated from the bile of various animals; several others have been prepared synthetically.

Lithocholic acid, 3-hydroxycholanolic acid, was first isolated by H. Fischer (65) from ox gall-stones; it is a normal constituent of ox bile and human bile. One gram could be isolated by Wieland (206) from 100 liters of ox bile which contained 5–6 kg. of cholic acid and 600–800 g. of deoxycholic acid. Lithocholic acid occurs presumably in a conjugated state like the other bile acids, although the actual isolation of its conjugated derivatives seems well-nigh impossible. It melts at 186°C.; its  $[\alpha]_D$  is about + 32°. It has a tendency to appear in gelatinous form. It may be separated from deoxycholic acid by its lesser acidity, as deoxycholic acid will migrate into an ethereal layer from a slightly acid aqueous solution, leaving the monohydroxy acid behind.

By far the most important of the dihydroxy acids is deoxycholic acid, 3,12-dihydroxycholanolic acid. Its similarity to the synthetic apocholic acid, 3,12-dihydroxycholenic acid, will be further discussed below (p. 360). When free of any retained solvent, it melts at 176.5°C. (corrected); its  $[\alpha]_D$  is + 55°. For its purification see Sobotka and Goldberg (175).

Isomeric dihydroxy acids are the 3,6-dihydroxy acid, hyodeoxycholic acid, m.p. 196–197°C., found in hog bile (85, 257, 262, 263), and the 3,7-dihydroxy acid, chenodeoxycholic acid (or anthropodeoxycholic acid), m.p. 108°C.,  $[\alpha]_D = +11^\circ$ , in goose bile (129, 259) and in human bile (217). 7,12-Dihydroxycholanolic acid, usually referred to as isodeoxycholic acid, is a synthetic product, melting at 226–227°C. (213).

3, 7,12-Trihydroxycholanolic acid, or cholic acid, is present in the bile of all animals. It melts at 195°C.; its  $[\alpha]_D = +37^\circ$ . Its isomer, 3,7,23-trihydroxycholanolic acid or  $\beta$ -phocaecholic acid, was found by Hammarsten in the bile of the seal (88); its constitution has been established by Windaus and Van Schoor (264). It melts at 22°C.; its  $[\alpha]_D = +27^\circ$  to +29°.

3,7,8,12-Tetrahydroxycholanolic acid (CXXI) (p. 341) was obtained by oxidation with permanganate from 3,12-dihydroxycholenic acid (26, 228). The natural occurrence of an isomeric tetrahydroxy acid is suspected, but not proven, in rabbit bile (264).

##### *B. Ketocholanolic acids and ketohydroxycholanolic acids*

The preparation and general reactions of the keto derivatives have already been discussed. The "dehydro derivatives" of all the hydroxy



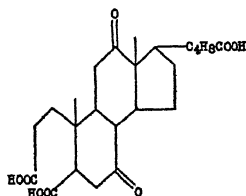
acids mentioned have been prepared, as well as partly oxidized derivatives containing both hydroxyl and keto groups. The 3-hydroxy-12-ketocholic acid, m.p. 160°C.,  $[\alpha]_D = +111^\circ$ , prepared by Cerecedo (33) from 3-monoacetyldeoxycholic acid by oxidation, occurs in human bile associated with an equivalent amount of chenodeoxycholic acid. This molecular compound, melting at 165°C., was first isolated by Weyland in Wieland's laboratory; its constitution was determined by Wieland and Kishi (234), who separated the hydroxyketo acid in the form of its semicarbazone.

### *C. Acids with three rings*

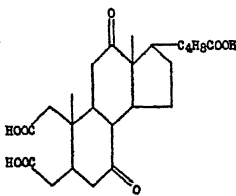
The simplest bile acid derivative containing three rings is the tricarboxylic acid, lithobilianic acid (LXXXIX), m.p. 275°C. (232), obtained by oxidation from dehydrolithocholic acid, where its formation is accompanied by that of isolithobilianic acid (XC). Other isomers obtained from the corresponding monoketo acids are thilobilianic acid (XCI) through cleavage of ring II between  $C_6$  and  $C_7$  (229) and another one (XCII) through cleavage of ring III between  $C_{11}$  and  $C_{12}$  (206, 209), all three isomers melting at about 260–262°C. Aetiobilianic acid, melting at 228°C. with opening of ring IV, after loss of the side chain, has already been mentioned (p. 317). Butenandt, Weidlich, and Thompson (31) recently obtained a dimethylphenanthrene, melting at 140°C., from aetiobilianic acid (XIV) by dehydrogenation and simultaneous double decarboxylation with selenium. An identical product was obtained from the hydrate of follicular hormone (CLIX), and its constitution as 1,2-dimethylphenanthrene was proved by Haworth's synthesis (91) (cf. p. 353).

Lithobilianic acid and isolithobilianic acid are the parent substances of a number of ketotricarboxylic acids derived from the various natural bile acids through their dehydro derivatives by further oxidation which attacks ring I whenever it contains oxygen on  $C_3$ . The acids obtained are bilianic acid (LXI), or 7,12-diketolithobilianic acid, m.p. 275°C.,  $[\alpha]_D = +76^\circ$ , and isobilianic acid (LXXXII), or 7,12-diketoisolithobilianic acid, m.p. 234–237°C. (115), substances whose isolation was one of the first achievements of the oxidative degradation of cholic acid. Deoxybilianic acid (LXXXIII), or 12-ketolithobilianic acid, m.p. 296°C.,  $[\alpha]_D = +94^\circ$  (204, 223, 225), was discovered likewise by Lachinov (116), who called this acid "cholic acid," a term abandoned later since it was applied to the mother substance of the bile acids (cf. 203). As well as its isomer, isodeoxybilianic acid (LXXXIV), Lachinov's "isocholic acid" of m.p. 247–248°C. and  $[\alpha]_D = +73^\circ$ , it is derived by oxidation from deoxycholic acid or by reduction, with zinc and fuming hydrochloric acid (21), of the more reactive

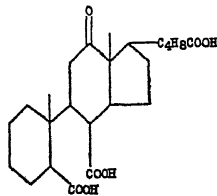
group  $C_7$  in bilianic and isobilianic acids, respectively. Chenodeoxybilianic acid (LXXXV), m.p.  $230^{\circ}\text{C}$ . (263), or 7-ketolithobilianic acid, is very easily reduced by Clemmensen's method to lithobilianic acid. Hyodeoxybilianic acid (LXXXVI), or 6-ketolithobilianic acid, has never been isolated on account of its  $\beta$ -keto group, which causes immediate decarboxylation to a ketodicarboxylic acid (LXXXVII). Hyoisodeoxybilianic acid exists only in its stereomeric or "allo"-form and will be discussed below (pp. 349, 350) under the name keto-Staden acid. The oxidation product of 7,12-diketochoLANic acid, dehydroisodeoxycholic acid, is known as pseudodeoxybilianic acid (LXXXVIII) m.p.  $259\text{--}260^{\circ}\text{C}$ . (21). Further oxidation of this acid leads to a pentacarboxylic acid discussed in the following section.



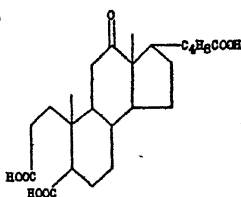
LXI  
Bilianic acid,  
 $\text{C}_{24}\text{H}_{34}\text{O}_8$



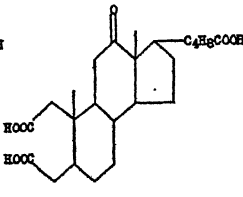
LXXXII  
Isobilianic acid,  
 $\text{C}_{24}\text{H}_{34}\text{O}_8$



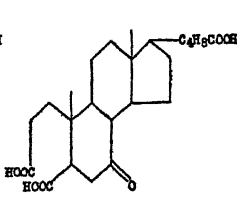
LXXXVIII  
Pseudodeoxybilianic acid,  
 $\text{C}_{24}\text{H}_{36}\text{O}_7$



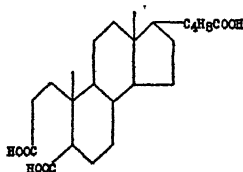
LXXXIII  
Deoxybilianic acid,  
 $\text{C}_{24}\text{H}_{36}\text{O}_7$



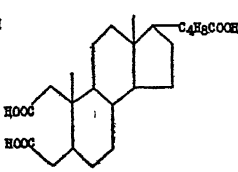
LXXXIV  
Isodeoxybilianic acid,  
 $\text{C}_{24}\text{H}_{36}\text{O}_7$



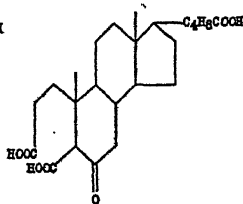
LXXXV  
Chenodeoxybilianic acid,  
 $\text{C}_{24}\text{H}_{36}\text{O}_7$



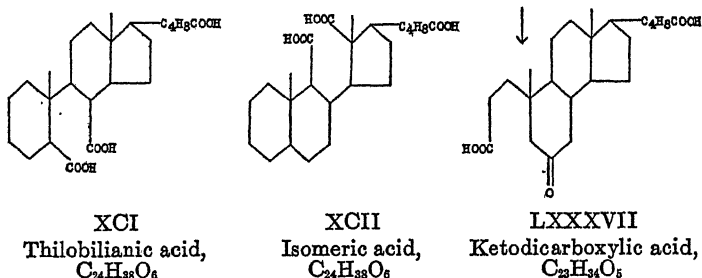
LXXXIX  
Lithobilianic acid,  
 $\text{C}_{24}\text{H}_{38}\text{O}_6$



XC  
Isolithobilianic acid,  
 $\text{C}_{24}\text{H}_{38}\text{O}_6$



LXXXVI  
(Hyodeoxybilianic acid,  
 $\text{C}_{24}\text{H}_{38}\text{O}_7$ )



Either oxidation of cholic acid with hypobromite or catalytic reduction of bilianic acid leads to 7-hydroxy-12-ketolithobilianic acid. The carboxyl group in  $C_4$  forms a  $\gamma$ -lactone with the hydroxyl group at  $C_7$ . This acid of melting point  $303-304^\circ\text{C}$ . (212, 22) was first prepared by Pringsheim (144), who gave it the name biliobanic acid (V) (p. 315). By further catalytic reduction, reductobiliobanic acid (XCIII), or 7,12-dihydroxylithobilianic acid, m.p.  $243^\circ\text{C}$ ., can be prepared (25). The isomeric isobiliobanic acid (XCV), melting at  $264^\circ\text{C}$ ., is the  $\epsilon$ -lactone of 12-hydroxy-7-ketoisolithobilianic acid. It cannot be prepared, however, from isobilianic acid (LXXXII, p. 335), as we know of no method that would reduce the carbonyl in  $C_{12}$  without reducing that in  $C_7$ . It can only be obtained by reduction of both keto groups in isobilianic acid to reductoisobiliobanic acid (LXXXIV), or 7,12-dihydroxyisolithobilianic acid, and subsequent partial oxidation of the  $>\text{CHOH}$  in position  $C_7$ .

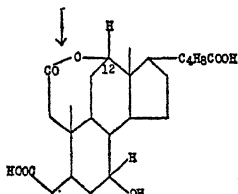
Another interesting substance of this type is chenodeoxybiliobanic acid (XCVI), obtained by Clemmensen reduction of the keto group at  $C_{12}$  in biliobanic acid, or by the action of hypobromite upon chenodeoxycholic acid (XCVII) (261, 222). The lactone melts at  $253^\circ\text{C}$ ., the free acid at  $255-256^\circ\text{C}$ . (22).

Thermic reclosure of ring I in the bilianic acid group leads to a number of pyrobilianic acids, all of which can be derived from pyrolithobilianic acid (XCVIII) and pyroisolithobilianic acid (XCIX). They contain four rings. Oxidative opening of this new cyclopentanone ring results in the formation of norbilianic acids, such as nordeoxybilianic acid (C), melting at  $232-233^\circ\text{C}$ . (210).

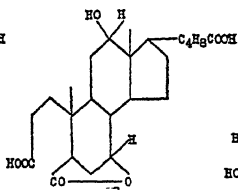
Oxidation of pyrodeoxybilianic acid (CI) with permanganate leads to a diketodicarboxylic acid (CII),  $C_{28}H_{34}O_6$ , m.p.  $216^\circ\text{C}$ . (204). This tricyclic acid, being relatively accessible, was made the starting point for several investigations by Wieland (215, 227). On further oxidation with permanganate, norcilianic acid (CIII) is obtained; a bicyclic diketotetracarboxylic acid melting at  $122^\circ\text{C}$ .

Oxidation of bilianic acid (LXI) with permanganate leads to a hypo-

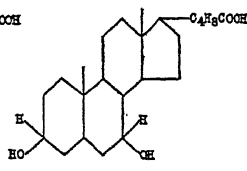
LXXXII  
Isobilianic acid,  
(see p. 335)



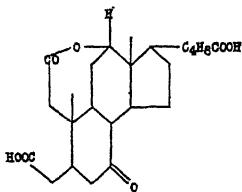
XCIV  
Redutoisobilianic  
acid,  
 $C_{24}H_{38}O_7$



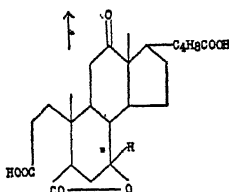
XCIII  
Redutobiliobanic  
acid,  
 $C_{24}H_{38}O_7$



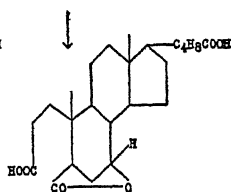
XCVII  
Chenodeoxycholic  
acid,  
 $C_{24}H_{40}O_4$



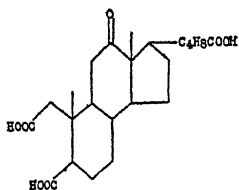
XCIV  
Isobilianic acid,  
 $C_{24}H_{38}O_7$



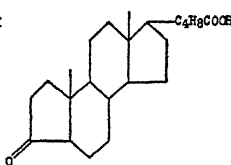
V  
Biliobanic acid, (Cheno)deoxybiliobanic acid,  
 $C_{24}H_{38}O_7$



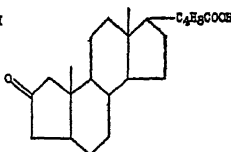
XCVI



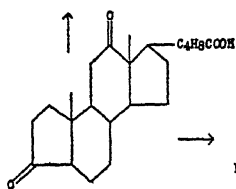
C  
Nordeoxybilianic  
acid,  
 $C_{23}H_{34}O_7$



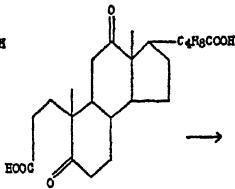
XCVIII  
Pyrolithobilianic  
acid,  
 $C_{23}H_{36}O_3$



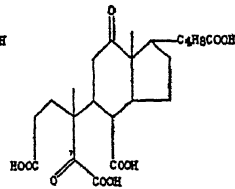
XCIX  
Pyroisolithobilianic  
acid,  
 $C_{23}H_{36}O_3$



CI  
Pyrodeoxybilianic  
acid,  
 $C_{23}H_{34}O_4$



CII  
Wieland's diketodicar-  
boxylic acid,  
 $C_{23}H_{34}O_6$



CIII  
Norcilianic  
acid,  
 $C_{23}H_{32}O_{10}$

thetical 6,7,12-triketolithobilianic acid (LXII), which suffers a benzilic acid rearrangement to cilianic acid (LXIII); this in turn, as mentioned on p. 330, is converted by permanganate and sulfuric acid into ciloxanic acid (LXIV), m.p. 216–217°C. (24).

Further oxidation of bilianic acids and their derivatives leads to bicyclic acids.

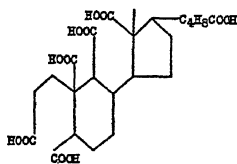
#### *D. Acids with two rings*

Just as lithobilianic acid is the simplest tricyclic bile acid derivative and may be considered, at least theoretically, as the mother substance of the bilianic acid group, so choloidanic acid is the prototype of bicyclic acids. It was first obtained by Cleve (36) from impure cholic acid, containing deoxycholic acid, by nitric acid oxidation. Lachinov (113) recognized deoxycholic acid as its only source (cf. reference 205), but was misled by certain similarities to camphoric acid and called the substance "chole-camphoric acid." It retains rings II and IV unaltered, while rings I and III have both been split to yield four carboxylic groups (XLVI) (see p. 327). M.p.  $\simeq$  300°C.;  $[\alpha]_D = +40^\circ$ .

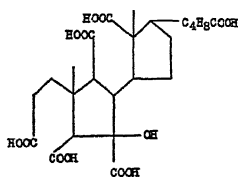
Closely related to choloidanic acid is chollepidanic acid (CIV), m.p. 280°C., which is obtained from deoxybilianic acid by strong oxidation with nitric acid. Its name is derived from the Greek word *λεπιδος*, *lepis*, meaning "scale," suggested by the appearance of its crystals. Its constitution is not completely elucidated, involvement of the side chain in the oxidative destruction may be ruled out, and Wieland assumes that the methyl group attached to  $C_{10}$  is oxidized to carboxyl. The acid is distinguished by its barium and calcium salts, which are less soluble in hot than in cold water (214, 223, 230).

Another acid in which rings I and III are disrupted is ciloidanic acid (CV), derived from cilianic acid (LXIII) by nitric acid oxidation. Its melting point is 248°C. (209). It bears the same relation to cilianic acid as choloidanic acid (XLVI, p. 327) does to deoxybilianic acid (LXXX, p. 335). In analogy with the transition of cilianic into ciloxanic acid (LXIV, p. 330), ciloidanic acid is converted into a ketopentacarboxylic acid (CVI) which, as a  $\beta$ -keto acid, is decarboxylated at once to a ketotetracarboxylic acid (CVII) melting at 238°C. (211).

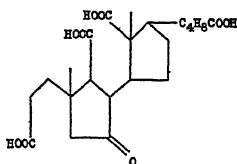
Thermic cyclization of ring I in choloidanic acid (XLVI, p. 327) yields pyrocholoidanic acid (CVIII). This acid has interesting properties, as it occurs in two stereomeric hydrated forms and as an unsaturated lactone. Inspection of a spatial model shows that lactonization probably takes place between the carboxyl in position  $C_{12}$  and the enolized carbonyl in  $C_3$  (CIX). Its oxidation with permanganate proceeds analogously to the oxidation of pyrodeoxybilianic acid to Wieland's diketodicarboxylic acid



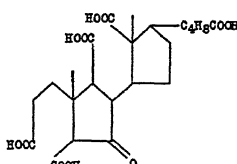
CIV  
Chollepidanic acid,  
 $C_{24}H_{34}O_{12}$



CV  
Ciloidanic acid,  
 $C_{24}H_{34}O_{13}$

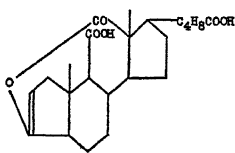


CVII  
Ketotetracarboxylic acid,  $C_{22}H_{32}O_9$

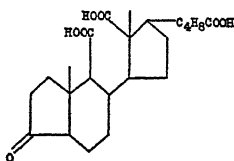


CVI

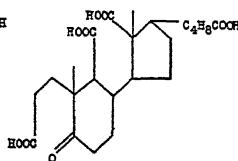
Ketopentacarboxylic acid,  
 $C_{23}H_{32}O_{11}$



CIX  
Lactone of CVIII,  
 $C_{23}H_{32}O_8$



CVIII  
Pyrocholoidanic acid,  
 $C_{23}H_{34}O_7$

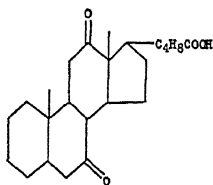


CX  
Prosolannelic acid,  
 $C_{23}H_{34}O_9$

(CI, CII, p. 337). The resulting ketotetracarboxylic acid, m.p.  $220^{\circ}\text{C}$ ., is designated prosolannelic acid (CX) for reasons cited below (207).

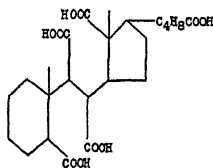
The number of bicyclic derivatives in which ring II is disrupted is much more limited. Rings II and III are opened in isocholoidanic acid (CXII), m.p.  $273^{\circ}\text{C}$ . (213), which is prepared from 7,12-diketocholanic acid (CXI) or the corresponding "pseudo" deoxybilanic acid (LXXXVIII) by treatment with nitric acid. Rings I and IV are intact and the carboxyl groups are in positions  $C_6$ ,  $C_7$ ,  $C_{11}$ , and  $C_{12}$  in addition to the original carboxyl in  $C_{24}$ .

A hydroxyl group was introduced by means of bromination and alkali into position  $C_{11}$  of 12-ketocholanic acid (CXIII), an isomer of dehydrolithocholic acid (231). Ring III was split by permanganate and the resulting keto acid (CXIV) was again hydroxylated on the tertiary carbon atom 8. The oxidation of this hydroxyketo acid to the ketotricarboxylic acid (CXV) of m.p.  $196^{\circ}\text{C}$ . is the first instance where the link  $C_8-C_9$  has



CXI

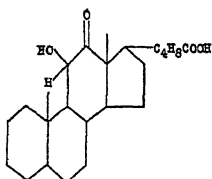
7,12-Diketochoholic acid,  
 $C_{24}H_{38}O_4$



CXII

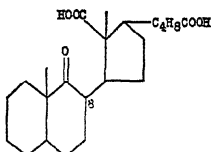
Isochoholic acid,  
 $C_{23}H_{36}O_{10}$

been disrupted while rings I and IV remain intact. Oxidation of this acid,  $C_{23}H_{36}O_7$ , with nitric acid (235) leads through simultaneous elimination of ten carbon atoms (Nos. 1-7, 9, 10, and 18) to the tricarboxylic acid,  $C_{13}H_{20}O_6$ , obtained previously by Wieland, Schlichting, and Vocke (216, 223).



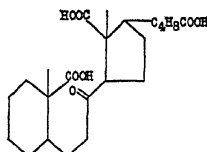
CXIII

11-Hydroxy-12-keto-  
choholic acid,  
 $C_{24}H_{38}O_4$



CXIV

Ketodicarboxylic acid,  
 $C_{23}H_{36}O_6$



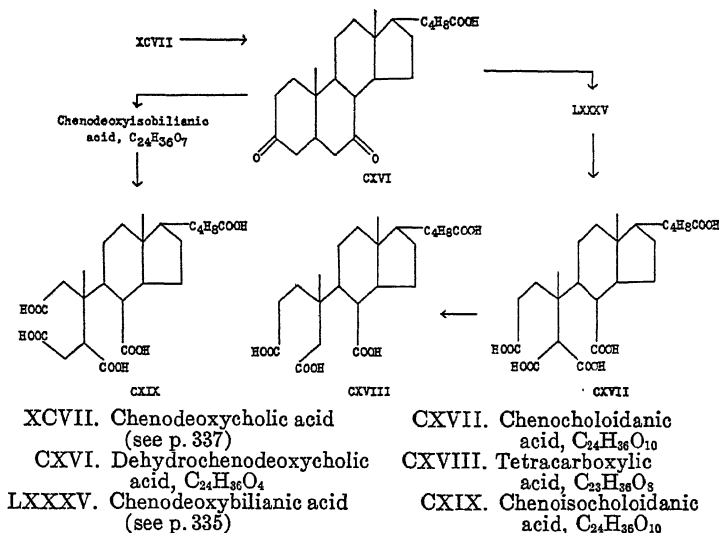
CXV

Ketotricarboxylic acid,  
 $C_{23}H_{36}O_7$

A third important isomer of cholidanic acid is chenochoholic acid (CXVII), m.p.  $225^{\circ}\text{C}$ ., obtained from the corresponding dihydroxy (XCVII), diketo- (CXVI), or bilianic (LXXXV) acid (218, 261). This acid with its carboxyl groups in positions 3, 4, 6, and 7 of the disrupted rings I and II, has two carboxyls attached to one carbon atom, i.e.,  $C_6$ ; because of this "malonic" character it is easily decarboxylated to a tetracarboxylic acid (CXVIII) of m.p.  $219-220^{\circ}\text{C}$ ., while its isomer, chenochoholic acid (CXIX), m.p.  $198-199^{\circ}\text{C}$ . (218), with carboxyl groups in positions 2, 3, 6, 7, and 24, is stable.

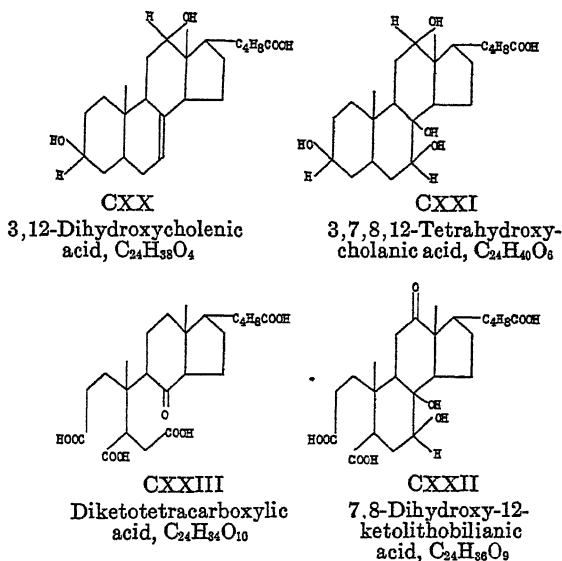
The oxidation of  $\Delta^{7,8}$ -3,12-dihydroxychoholic acid (CXX) leads by way of 3,7,8,12-tetrahydroxychoholic acid (CXXI, (26) see p. 333) and 7,8-dihydroxy-12-ketolithobilianic acid (CXXII) to a diketotetracarboxylic acid (CXXIII), whose properties and constitution corroborate the assumption of the double bond between  $C_7$  and  $C_8$  in dihydroxychoholic acid (228, 233).

Bicyclic derivatives with rings IV and III open seem to be formed by oxidative destruction from the female sex hormone (124).



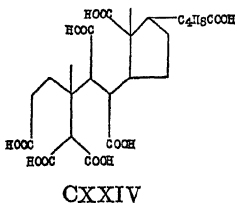
### E. Acids with one ring

The scission of the three rings I, II, and III, would yield a heptacarboxylic acid of the formula CXXIV by analogy with lithobilianic acid,



choloidanic acid, or chenocholoidanic acid. The close interrelationship between rings I and II, however, is not in favor of the accumulation of so



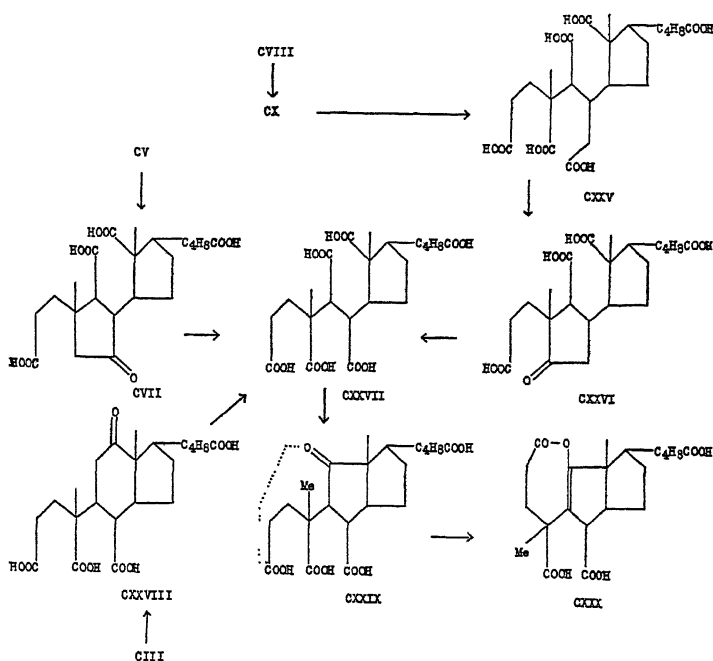


many carboxyl groups and, as a result, no monocyclic oxidation products containing the complete twenty-four carbon atom skeleton is known. Two lower homologs take its part in the scheme of oxidative degradation of bile acids: biloidanic acid (CXXVII),  $C_{22}H_{32}O_{12}$ , m.p. 226–228°C., first isolated by Letsche (119), and Wieland's solannellic acid (CXXV) (*solus annellus* = one ring only),  $C_{23}H_{34}O_{12}$ , m.p. 202–203°C. (207).

The latter acid is formed by oxidative scission of ring II, between the carbonyl in  $C_5$  and the methylene in  $C_6$  of prosolannellic acid (CX). This acid has been previously mentioned (p. 339), as an oxidation product of pyrocholoidanic acid. Solannellic acid was also obtained directly from pyrocholoidanic acid (CVIII). Ring II, which has been rebuilt as a cyclopentanone ring in pyrosolannellic acid, m.p. 272–273°C. (207), is reopened by nitric acid. The hexabasic acid formed was called norsolannellic acid by Wieland (211), but he soon established its identity with the well-known biloidanic acid, first isolated by Letsche and named by Schenck (166a). This hexabasic acid with twenty-two carbon atoms was prepared by the action of nitrating mixture ( $HNO_3 + H_2SO_4$ ) upon cholic acid (119) and by the action of fuming nitric acid on cholic acid, bilianic acid, or pyrodeoxybilianic acid (CI) (207, 208). More perspicuous reactions resulting in the formation of biloidanic acid are (1) the oxidation of the ketotetracarboxylic acid (CVII),  $C_{22}H_{32}O_9$ , from ciloidanic acid (CV, p. 339), (2) the oxidation of the isomeric ketotetracarboxylic acid (CXXVIII), derived from norcilianic acid (CIII, p. 337) through loss of carbon monoxide, and (3) the oxidation of a third isomeric ketotetracarboxylic acid, namely, pyrosolannellic acid (CXXVI) as mentioned above.

Both solannellic acid and biloidanic acid are quite stable against further oxidation; the remnants of the three rings are disposed in such a manner that carboxyl groups are separated by at least two carbon atoms.

While thermic cyclization of solannellic acid leads to a cyclopentanone in place of ring II, biloidanic acid under "pyro" conditions fuses the carboxyl groups  $C_{11}$  and  $C_{12}$  into a keto group (CXXIX), which in turn forms an unsaturated lactone with carboxyl group  $C_3$  (CXXX). The cyclopentanone structure of ring III in pyrobiloidanic acid is noteworthy,



- CIII. Norcilianic acid (p. 337)  
 CV. Ciloidanic acid (p. 339)  
 CVII. Ketotetracarboxylic acid (p. 339)  
 CVIII. Pyrocholoidanic acid (p. 339)  
 CX. Prosolanelllic acid (p. 339)  
 CXXV. Solannellic acid,  $C_{23}H_{34}O_{12}$

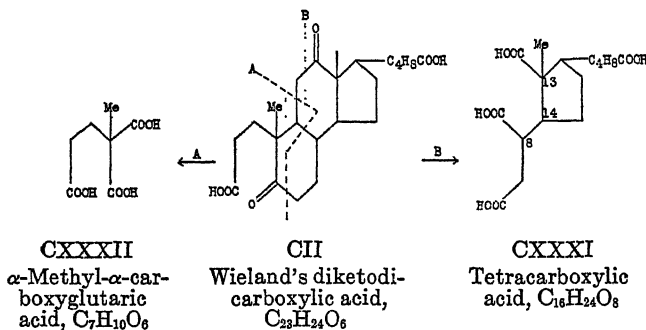
- CXXVI. Pyrosolanelllic acid,  $C_{23}H_{32}O_9$   
 CXXVII. Biloidanic acid (= nor-solanelllic acid),  $C_{22}H_{32}O_{12}$   
 CXXVIII. Ketotetracarboxylic acid,  $C_{22}H_{32}O_9$   
 CXXIX. Pyrobiloidanic acid,  $C_{21}H_{30}O_9$   
 CXXX. Unsaturated lactone of CXXIX,  $C_{21}H_{28}O_8$

since all other derivatives with carboxyl groups in  $C_{11}$  and  $C_{12}$  fail to behave as expected of 1,6-dicarboxylic acids, according to "Blanc's rule."

The various possibilities of partial and complete esterification in these polybasic acids deserve mention; they may be studied in the original publications, e.g., in the paper by Wieland and Schlichting on the various esters of biloidanic acid (208).

While most methods of oxidative degradation seem to be exhausted at this stage, oxidation of Wieland's diketodicarboxylic acid,  $C_{23}H_{38}O_6$ , leads further to the important tetracarboxylic acid (CXXXI),  $C_{16}H_{24}O_8$ , of m.p.  $218^\circ\text{C}$ . (216, 223, 224). This acid was first obtained as a by-product in the oxidation of pyrodeoxybilianic acid to Wieland's diketo acid. Better yields can be obtained by the use of nitrating mixture on the latter acid.

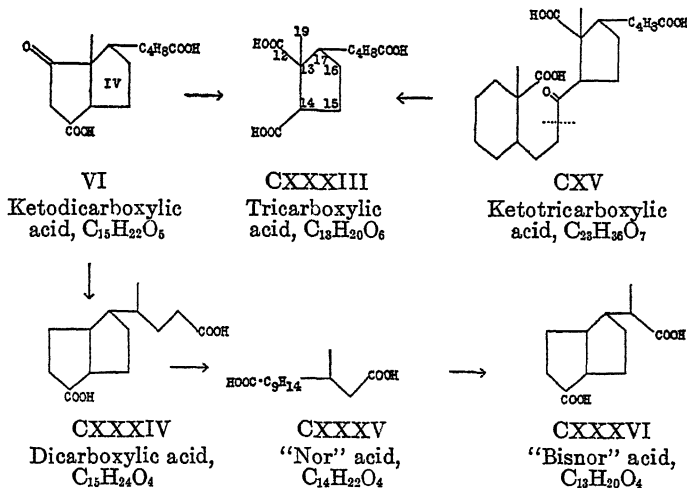
Yields of 14.5 per cent by weight of the starting material, deoxycholic acid, were reached by Wieland. The carboxyl group  $C_{12}$  behaves rather passively and this indicates the presence of a methyl group  $C_{19}$  attached to  $C_{13}$ . This assumption is further supported by the formation of acetone upon reaction with hydrogen peroxide, which finds a parallel in the case of  $\alpha$ -methylglutaric acid or  $\alpha$ -methyladipic acid (cf. p. 318). By-products in the preparation of this acid are succinic acid and  $\alpha$ -methyl- $\alpha$ -carboxyglutaric acid (CXXXII) (= *n*-butane-1,3,3'-tricarboxylic acid (223, 192)). The simplest explanation of the appearance of these acids seems to be that indicated by the broken lines in formula CII in agreement with the assumption of a methyl group ( $C_{18}$ ) on  $C_{10}$  (cf. reference 156). A green intermediary nitroso derivative indicates that the point of attack for the nitric acid (or rather for the nitrous acid) is the tertiary carbon atom  $C_9$ .



Thermic cyclization of the tetracarboxylic acid  $C_{16}H_{24}O_8$  (CXXXI) leads to the ketodicarboxylic acid (VI),  $C_{15}H_{22}O_6$ , m.p.  $187^\circ\text{C}$ ., in poor yields (7 g. from 145 g.). This pyro acid can be deprived of two more carbon atoms in two different series of reactions as follows: (a) Nitric acid or bromine (cf. p. 316) disrupt the cyclopentanone ring, and both lead to tricarboxylic acid (CXXXIII),  $C_{13}H_{20}O_6$ , also melting at  $187^\circ\text{C}$ ., which was recently obtained by another route, namely from acid CXV (235). This acid forms an anhydride, and the carboxyl group on  $C_{13}$  displays the same inactivity as the analogous carboxyl group in  $C_{16}H_{24}O_8$ . (b) Reduction by Clemmensen's method to  $C_{15}H_{24}O_4$  (CXXXIV) furnished a starting material for a Grignard reaction along the lines exemplified in the case of cholanic acid  $\rightarrow$  norcholanic acid  $\rightarrow$  bisnorcholanic acid. The carboxyl groups  $C_9$  and  $C_{24}$  are converted into diphenylcarbinol groups. Whilst the former remains protected against oxidation,  $C_{24}$  and subsequently  $C_{23}$  can be eliminated (CXXXV, CXXXVI). Further progress was impossible because of the small amounts of material available (224).

The formation of the substances mentioned in the preceding paragraph

is in complete agreement with all other evidence and confirms decisively the conceptions developed during the last year. The five-membered nature of ring IV has been corroborated by the formation of dimethylphenanthrene from aetiobilianic acid (see pp. 334 and 353).



Conclusive evidence for the allocation of the methyl group  $C_{19}$  to  $C_{13}$  (and not  $C_{14}$ ) lies in the steric configuration of the cyclopentanone ring in the pyro acid  $C_{15}H_{22}O_5$ . Its cleavage product, the tricarboxylic acid  $C_{13}H_{20}O_6$  (CXXXIII), m.p.  $187^\circ C$ ., bears the two carboxyl groups on  $C_{13}$  and  $C_{14}$  in the *trans*-position, since a lower melting, more soluble isomer (m.p.  $137^\circ C$ .) is obtained via the common anhydride. Thus the cyclopentanone ring in  $C_{15}H_{22}O_5$ , too, must have been attached in the *trans*-position. Ordinarily, such cycloketones would assume the *cis*-position, if necessary by an inversion. This rearrangement is however impossible in the parent substance  $C_{16}H_{22}O_8$  (CXXXI, on pp. 343 to 344), if we assume that  $C_{13}$  carries no H, but the methyl group  $C_{19}$  (235).

Finally, the side chain can now be definitely allocated to  $C_{17}$  (instead of the alternative  $C_{15}$ ), both because of consideration of molecular dimension and shape, and because of ring closures between carbonyl  $C_{12}$  and methylene  $C_{23}$  of the side chain, an example of which is given on p. 329 (156). The side chain is in the *cis*-position to the link  $C_{12}-C_{13}$ , i.e., in the *trans*-position relative to the methyl group,  $C_{19}$  (235).

#### F. Steric isomerisms

The study of the monocyclic derivatives discussed in the preceding section throws some light on the relative steric configuration of the asym-

metric carbon atoms  $C_{13}$ ,  $C_{14}$ , and  $C_{17}$ . Similar considerations, e.g., of the pyro reactions of dihydro-Diels acid and its isomer (pp. 349, 350), let us assume that the cholan acid-coprostan series bears the same relation to the allocholan-cholestan series in regard to the relative position of rings I and II as *cis*-decalin does to *trans*-decalin. The steric relation between rings II and III is as yet unknown.

There is no more doubt, however, that the cyclic portion of the carbon skeleton is identical in these two large biological groups except for the steric configuration of one carbon atom, which can be only  $C_5$ , according to the formula of Rosenheim and King.

A glance at the formula of cholan acid reveals eight asymmetric carbon atoms, viz.,  $C_5$  and  $C_{10}$ ,  $C_8$  and  $C_9$ ,  $C_{13}$  and  $C_{14}$ ,  $C_{17}$  and  $C_{20}$ . Even on the assumption that certain steric isomers are possible in theory only, but could not be realized because of excessive distortion of the ring system, it seems remarkable that sterols and bile acids should tally except for the constellation on a single carbon atom. This peculiarity reminds one of the unique prevalence of *d*-glucose among the numerous possible stereomeric aldohexoses.

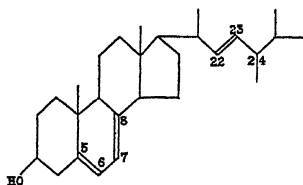
As laws of mutual induction seem to govern biosynthesis of the carbon skeleton, the position of the hydroxyl groups is likewise subject to a natural selection. While the secondary alcohol groups in positions  $C_3$ ,  $C_6$ ,  $C_7$  and  $C_{12}$  create additional centers of asymmetry and offer opportunities for "epimerism," only one of the alternatives is ever realized in natural products. A few epimers have been artificially prepared in the sterol series by inversion of the —OH on  $C_3$ . Treatment with sodium in amyl alcohol catalyzes the establishment of equilibria between cholestanol, m.p. 141°C. (dihydrocholesterol, also  $\beta$ -cholestanol), and epicholestanol, m.p. 182°C. ( $\epsilon$ -cholestanol (249)). Analogously, coprostanol (stercorin (66), coprosterol (20) ) m.p. 99–100°C., had been converted in good yield into epicoprostanol, melting at 119°C., by Dorée and Gardner ("pseudo-coprosterol" (55, 56), also  $\delta$ -cholestanol (250, 251) ). These epimers are distinguished from the natural sterols by their inability to form insoluble addition compounds with saponins (cf. p. 359).

#### *G. Sterols with longer side chains*

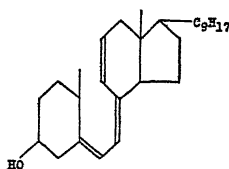
Ergosterol (CXXXVII),  $C_{28}H_{48}OH$ , m.p. 160°C., discovered by Tanret (182) in ergot, occurs in a variety of fungi, including yeast, where it is accompanied by zymosterol, m.p. 109°C. (122, 93, 149). Since ergosterol was recognized to be provitamin D, its chemistry has been studied extensively. It contains the same carbon skeleton as cholesterol, with an additional methyl group on  $C_{24}$ . Of its three double bonds, one is located in the side chain between  $C_{22}$  and  $C_{23}$  (formation of methylisopropylacetal-

dehyde with ozone (150, 82), see p. 319), the other two are in conjugation with each other in ring II ( $C_5=C_6$ ,  $C_7=C_8$ ) (159).

By ultra-violet irradiation of ergosterol, a number of isomers are formed, all of which respond to the formula of ergosterol, but none of which is precipitable with digitonin; they carry the names lumisterol, protachysterol (?), tachysterol, calciferol (crystallized vitamin D), toxisterol (?), supra-sterol I and II. Only vitamin D has antirachitic properties. Lettré (269) has shown that tachysterol contains four double bonds, i.e., one more than ergosterol, and suggests that irradiation causes cleavage of ring II and simultaneous formation of a new double bond, by analogy with the photochemically induced cleavage of the cyclohexanone ring (XLV) to hexenaldehyde  $\text{CH}_2=\text{CH}\cdot\text{CH}_2\cdot\text{CH}_2\cdot\text{CH}_2\cdot\text{CHO}$  (270). Formula CXXXVIII for tachysterol would explain the peculiarities of the irradiation products mentioned, their inability to combine with digitonin, and their failure to yield tetracyclic hydrocarbons on selenium dehydrogenation. Migration of the double bonds upon further irradiation is perhaps the mechanism underlying the formation of vitamin D and subsequently of the end products in which the conjugated system of double bonds no longer exists, thus abolishing further absorption of and sensitivity against ultra-violet



CXXXVII  
Ergosterol,  $\text{C}_{28}\text{H}_{48}\text{OH}$



CXXXVIII  
Tachysterol,  $\text{C}_{28}\text{H}_{48}\text{OH}$   
(Lettré, 269)

rays. A number of other ergosterol isomers have been synthesized by purely chemical methods, but none of them could be identified with any of the irradiation products, nor did they display any vitamin D activity.

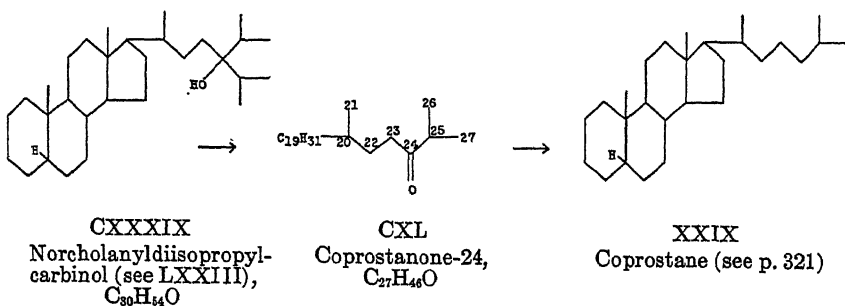
The best known representatives of the phytosterols, the sterols of higher plants, are sitosterol  $\text{C}_{29}\text{H}_{49}\text{OH}$ , m.p.  $137^\circ\text{C}$ . (*σιτος*, Greek, "wheat") and stigmasterol,  $\text{C}_{29}\text{H}_{47}\text{OH}$ , m.p.  $170^\circ\text{C}$ . (*Physostigma venenosum* = calabar bean). One of the two double bonds of the latter is located in the side chain as in ergosterol, but this sterol carries an *ethyl* group on  $\text{C}_{24}$ . The only double bond of sitosterol can be hydrogenated; the resulting saturated sterol, sitostanol, has likewise been found in nature. It is not identical, but is probably stereomeric, with stigmastanol, the tetrahydro derivative of stigmasterol (258, 6, 165, 64).

It may be mentioned in passing that certain saponins (amyrin, rhamnol, etc.) and cardiac poisons, as strophanthidin (101), resemble the sterols structurally and may be classified with them. Phytosterolins, glucosides of phytosterols, are well-defined substances and some have been synthesized.

Certain lower animals contain sterols different from cholesterol; spongo-sterol from sponges, asteriasterol from star fish eggs (140), bombycestrol from silk worms and several others have not been proven to be chemical individuals, but the sterol of the oyster and other bivalves has recently been investigated (12). This ostreasterol melting at  $140^{\circ}\text{C}$ . is an isomer of stigmasterol,  $\text{C}_{29}\text{H}_{47}\text{OH}$ ; hydrogenation of its two double bonds leads to sitostanol. Wool fat (lanolin) contains two sterols, lanosterol,  $\text{C}_{30}\text{H}_{48}\text{OH}$ , m.p.  $141^{\circ}\text{C}$ . and agnosterol,  $\text{C}_{30}\text{H}_{47}\text{OH}$ , m.p.  $162^{\circ}\text{C}$ .; they are not precipitable with digitonin, which increases the difficulty in the study of these interesting cutaneous excretion products (266).

#### *H. Transitions. Modifications of the side chain*

The structural identity or near-identity of sterols and bile acids was first verified experimentally by Windaus and Neukirchen (252), when they succeeded in the conversion of cholestane, by oxidation of the terminal group, into "allocholanic acid," and the conversion of coprostanane ("pseudo-cholestane") into regular cholanic acid. Similarly, 3-chlorocholestane



was converted into allolithocholic acid (260). Later, Wieland and Jacobi (238), reversed this process. Ethyl cholanate was converted by Grignard's reagent into the diisopropylcarbinol (CXXXIX) (cf. p. 331), which, in turn, on oxidation with chromium trioxide gave the isopropyl norcholanyl ketone (cholanyl- $\beta$ -propane). This substance is nothing else than coprostanone-24 (CXL), and may be reduced to coprostanane (XXIX) by zinc amalgam.

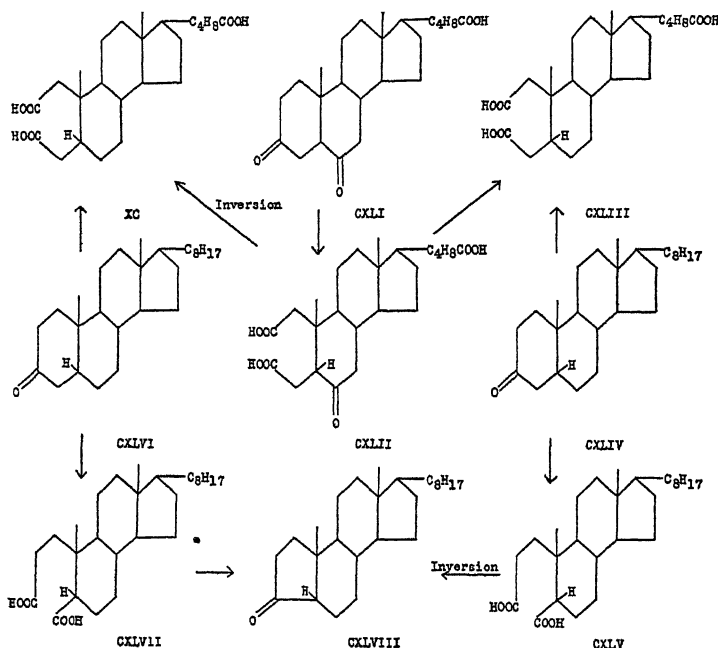
These interconversions between the sterol group and the bile acids have

been supplemented by a number of subsequent investigations. In some cases, inversion of the substituents on  $C_5$  allows the preparation of both isomers from a common starting material, e.g., cholestene can be converted via its hydrochloride into coprostene, and while cholestene itself is hydrogenated to cholestane, coprostene is hydrogenated either to coprostane (= pseudocholestane) or to cholestane, depending on the procedure (see p. 320). The converse, where an identical product is attained from stereomeric sources, will be exemplified by three instances. (a) Hyodeoxycholic acid, itself a cholanic acid derivative, is dehydrogenated to a labile  $\alpha$ -dehydrohyodeoxycholic acid, m.p.  $162^\circ\text{C}$ .,  $[\alpha]_D = -66^\circ$ , which suffers a Walden inversion on  $C_5$  owing to the vicinity of the carbonyl group. (Intermediary enolization?) Oxidative cleavage of ring I in the stable  $\beta$ -dehydro acid (CXLI) of m.p.  $209\text{--}210^\circ\text{C}$ .,  $[\alpha]_D = 0$  (?), leads to 6-keto-isolithobilianic acid (CXLII), which is reduced by Kishner and Wolff's method to a mixture of two isolithobilianic acids. One of them is identical with Staden's acid (CXLIII) obtained from dihydrocholesterol via cholestanone (CXLIV) (255), the other one with isolithobilianic acid (XC) obtained from dehydrolithocholic acid and also from coprostanol via coprostanone (CXLVI). (b) Cholestanone, a saturated ketone, is oxidized to  $\text{C}_{27}\text{H}_{46}\text{O}_4$ , the dihydro derivative of Diels' acid (CXLV) (see p. 316). Coprostanone can be oxidized to a stereomeric dicarboxylic acid (CXLVII). Both acids yield the same pyroketone (CXLVIII) by thermic cyclization, which indicates, according to Windaus, that the center of asymmetry responsible for the isomerism of the acids must be vicinal to one of the carboxyl groups, i.e., in position 5.

The reaction runs more smoothly and with better yield in the case of the "copro" isomer. As such cyclic pyroketones usually assume *cis*-configuration relative to the adjacent nucleus, this difference speaks for *cis*-configuration in the copro series, while the dihydro-Diels acid from cholesterol had to undergo an inversion of the carboxyl on  $C_5$  before formation of the *cis*-ketone. (c) Wieland (231) carried out the following series of reactions: pyrolithobilianic acid (CXLIX) was oxidized by permanganate to 5-keto-3,24-dicarboxylic acid (CL), this in turn by hypobromite to the 3,5,6,24-tetracarboxylic acid (CLI),  $\text{C}_{23}\text{H}_{36}\text{O}_8$ . Its pyro derivative (CLII),  $\text{C}_{22}\text{H}_{34}\text{O}_5$ , is oxidized by nitric acid to the 3,5,7,24-tetracarboxylic acid (CLIII),  $\text{C}_{22}\text{H}_{34}\text{O}_8$ . This acid, m.p.  $194^\circ\text{C}$ ., was identical with a tetrabasic acid obtained by Windaus (247) in the cholesterol series from Diels' acid (formula VIII). The last instance demonstrates the convergency of the two stereomeric series into a common end product through disappearance of the asymmetry of carbon atom  $C_5$ .

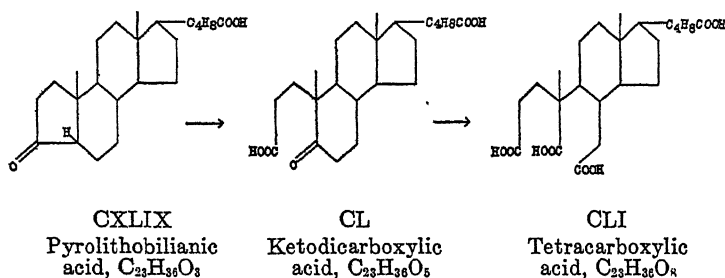
These synthetic links between the sterol series and the  $\text{C}_{24}$ -acid series series have found parallels in nature. Wieland and Kishi (234), were able

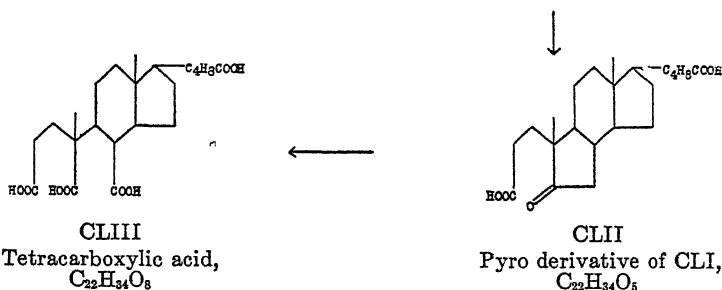




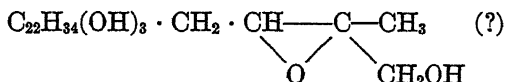
- XC. Isolithobilianic acid (see p. 335)  
 CXLI.  $\beta$ -Dehydrohyodeoxycholic acid,  $C_{24}H_{38}O_4$   
 CXLII. 6-Ketoalloisolithobilianic acid (= hyodeoxyisobilianic acid),  $C_{24}H_{38}O_7$   
 CXLIII. Allo isomer of XC (Staden's acid),  $C_{24}H_{38}O_5$   
 CXLIV. Cholestanone,  $C_{27}H_{48}O$   
 CXLV. Dihydro-Diels' acid,  $C_{27}H_{46}O_4$   
 CXLVI. Coprostanone,  $C_{27}H_{46}O$   
 CXLVII. Copro isomer of dihydro-Diels' acid,  $C_{27}H_{46}O_4$   
 CXLVIII. Dihydropyro-Diels' ketone,  $C_{26}H_{44}O$

to isolate from the mother liquors of 10 tons of ox bile a few grams of a new acid, which is distinguished from all other bile acids by the number of carbon atoms.  $C_{28}H_{46}O_4$  was given as a preliminary formula of this "sterocholic acid," a name suggesting a closer relation to the sterols. Other products which might fill the place of the missing link between sterols





and bile acids are tetrahydroxybufostane, C<sub>27</sub>H<sub>44</sub>(OH)<sub>4</sub> (?) found in toad bile (125), and especially scymnol. Although all vertebrate animals produce in their liver a biliary fluid, no bile acid in the proper sense is found in the bile of one of the phylogenetically lowest vertebrates, the shark (*scymnus borealis*). Scymnolsulfuric acid, discovered by Hammarsten (87a) in shark bile, is a sulfuric ester of scymnol, C<sub>27</sub>H<sub>46</sub>O<sub>8</sub>, m.p. 187°C. Three of its oxygen atoms are apportioned to three secondary alcohol groups on C<sub>7</sub>, C<sub>12</sub> and in ring I, and the fourth one to a primary alcohol group on C<sub>27</sub>, whilst the fifth is an ethylene oxide oxygen, forming a three-atom ring with C<sub>24</sub> and C<sub>25</sub> (184, 237, 240, 241, 265):



Certain toad poisons, studied by Wieland, by several Japanese investigators (109), and especially by K. K. Chen, H. Jensen, and A. L. Chen (271), are structurally related to the bile acids (see p. 313). No complete structural formulas have been established for them, owing to the scarcity of the starting material.

Rosenheim and Starling (160) recently obtained the elusive "oxycholesterol" (C<sub>26</sub>H<sub>46</sub>O<sub>2</sub>, m.p. 236°C.) of Lifschütz (121) in the crystalline state by oxidation of cholesterol in various ways. An isomer melting at 176°C. was obtained (32) by the action of selenium dioxide on cholesterol. Both glycols, probably the *trans*- and *cis*-forms of 3,4-dihydroxycholestene (Δ<sup>5,6</sup>), rearrange easily into *coprostenone*, 3-ketocoprostone (Δ<sup>4,5</sup>). This unsaturated ketone yields on reduction *coprostenol* (allocholesterol). The latter is (a) easily reconverted into cholesterol, which hinders its detection as an intermediary in biological processes; (b) it can, in turn, be hydrogenated to *coprostanol* (coprosterol) which has never been obtained *in vitro* by direct reduction of cholesterol, but seems to be an end product of sterol metabolism; (c) coprostenol possesses the "copro" configuration and may thus be the biological predecessor of the bile acids.

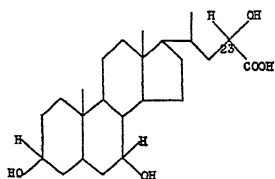
*J. Degradation of side chain; hormones and carcinogenic products*

Nothing is known in regard to the biological destruction of the bile acids. A bile acid with modified side chain is  $\beta$ -phocaecholic acid (CLIV), the 3,7,23-trihydroxycholanolic acid referred to on p. 333. The typical hydroxyl group on  $C_{23}$  is located incidentally on one of the atoms which carry a double bond in the side chain of some of the sterols. Permanganate converts the bile acid of the seal into  $C_{23}H_{38}O_4$ , which is presumably 3,7-dihydroxynorcholanolic acid.

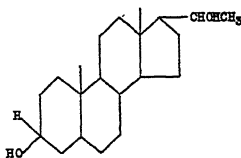
Pregnandiol (CLV),  $C_{21}H_{36}O_2$ , m.p.  $236^\circ\text{C}$ ., a di-secondary alcohol recovered as a by-product in the isolation of female sex hormone from pregnancy urine is likewise a derivative of cholanolic acid (30). One of its hydroxyl groups is most probably in position  $C_3$  in accordance with all other natural products dealt with in this review; the other one belongs to the short side chain. The diol is oxidizable to a diketone, which in turn is reduced to the tetracyclic hydrocarbon pregnane (CLVI),  $C_{21}H_{36}$ , m.p.  $83.5^\circ\text{C}$ . This substance was also obtained by reduction of aetiocholanyl methyl ketone (CLVII), m.p.  $115^\circ\text{C}$ ., prepared from bisnorcholanolic acid methyl ester via the diphenylcarbinol and dehydration to the diphenylethylene (LXXVIII) (p. 332) and, finally, oxidation of the latter by ozone. While pregnandiol is physiologically inactive (71), a closely related unsaturated diketone  $C_{21}H_{30}O_2$  (CLXIII), appearing in two polymorphous modifications melting at  $120^\circ\text{C}$ . and  $128^\circ\text{C}$ ., has been recently identified as the active component of crystalline corpus luteum concentrates. (Isolation of crystals (272); structure (273); conversion of pregnandiol into "luteosterone" (274); preparation from stigmasterol (275).)

Removal of the last two carbon atoms ( $C_{21}$  and  $C_{20}$ ) of the side chain leads to the male and female sex hormones. The male hormone, androsterone (CLXIV), m.p.  $178^\circ\text{C}$ ., is a saturated hydroxyketone,  $C_{19}H_{30}O_2$ . Its structure, 3-hydroxyaetioallocholanone-17, has been made probable by Butenandt (276), and corroborated by Ruzicka (277), who synthesized it by oxidation from epidihydrocholesterol ( $\epsilon$ -cholestanol) (cf. p. 346).

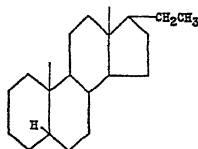
The follicular hormone  $C_{18}H_{22}O_2$  (ketohydroxyoestrin = theelin) and its hydrate  $C_{18}H_{24}O_3$  (trihydroxyoestrin = theelol) have been given the formulas CLVIII and CLIX. The older literature on isolation and properties of these substances was summarized recently by Marrian (128) and by Girard (75). For nomenclature see reference 3. These substances are the 3-hydroxy-17-keto and the 3,16,17-trihydroxy derivatives (126) of the hypothetical hydrocarbon oestratriene (CLX),  $C_{18}H_{24}$ , which is cholane minus the side chain  $C_{20}$  to  $C_{24}$  and the methyl group  $C_{10}$  and with three aromatic double bonds in ring I. Marrian and Haslewood (127) and McCorquodale, Thayer, and Doisy (123) opened ring IV of the trihydroxy hormone by alkali fusion.



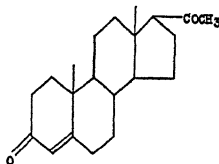
CLIV  
 $\beta$ -Phocaecholic acid,  
 $C_{24}H_{40}O_5$



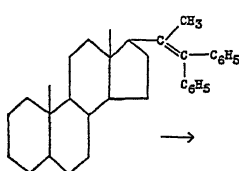
CLV  
Pregnandiol,  
 $C_{24}H_{38}O_2$



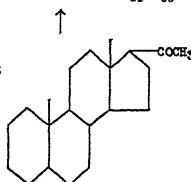
CLVI  
Pregnane,  
 $C_{21}H_{36}$



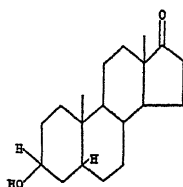
CLXIII  
Corpus luteum  
hormone,  
 $C_{21}H_{30}O_2$



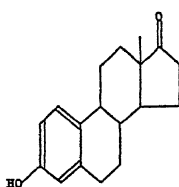
LXXVIII  
Diphenylbisonorcholene,  
(see p. 332)



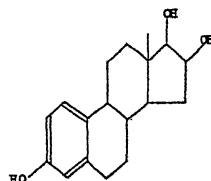
CLVII  
Aetiocolanlyl  
methyl ketone  
(= pregnanone-  
20),  $C_{21}H_{34}O$



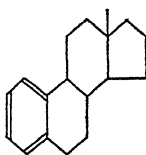
CLXIV  
Androsterone,  
 $C_{19}H_{30}O_2$



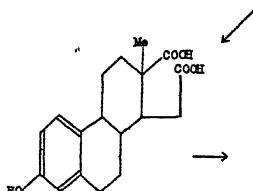
CLVIII  
Ketohydroxyoestrin,  
 $C_{18}H_{32}O_2$



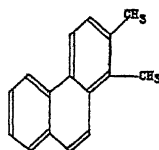
CLIX  
Trihydroxyoestrin,  
 $C_{18}H_{24}O_3$



CLX  
Oestratriene,  
 $C_{18}H_{24}$



CLXI  
Phenolic dicarboxylic  
acid,  $C_{18}H_{22}O_5$



CLXII  
Dimethylphenanthrene,  
 $C_{14}H_{14}$

The resultant phenolic dicarboxylic acid (CLXI) possesses the same skeleton as aetiobilianic acid (XIV). Complete dehydrogenation and simultaneous decarboxylation lead, as with aetiobilianic acid, to dimethylphenanthrol and subsequent distillation with zinc dust to the 1,2-dimethylphenanthrene (CLXII) mentioned on p. 334. The position of the hydroxyl group on  $C_3$  (cholane nomenclature = position 7 in 1,2-di-

methylphenanthrene) is extremely probable, but awaits direct verification. This work establishes the close configurational and presumably genetic relationship between the sex hormones and the cholane and cholestane group.

The hydrocarbons obtained by "aromatization" suggest a connection between the compounds reviewed in these pages and the synthetic hydrocarbons of the London Cancer Hospital research group (38, 39, 40, 105, also 279). Their polycyclic aromatic hydrocarbons, usually embodying a phenanthroid constellation, display carcinogenic, and often also oestrogenic, activity, and are held responsible for the experimental generation of tumors by coal tar. Synthetic derivatives such as 1,2,5,6-dibenanthracene belong here as well as methylcholanthrene (XXIVA) obtained from bile acid derivatives by aromatization. As "the cell proliferation which characterizes the oestrous state is in some respects reminiscent of the early stages of malignant growth" (quoted from reference 39), one may suspect that dehydrogenation of sterols and bile acids is the physiological method for the production of the sex hormone, but that further dehydrogenation by a faulty mechanism gives rise to carcinogenic substances (105, 128).

#### V. CHEMICAL AND PHYSICAL PROPERTIES OF BILE ACIDS

A list of all bile acid and sterol derivatives known to date would run into a total sum of five hundred and more, even if all functional derivatives such as esters, oximes, lactones, and halogen substitution products were disregarded. The biochemistry of the degradation products of proteins, carbohydrates, and lipoids has been intensely studied, since the products obtained by hydrolytic decomposition of these big molecules are identical with the intermediaries in animal and plant metabolism, e.g., amino acids, purine bases, monosaccharides, fatty acids, and so on. On the other hand, the numerous derivatives of bile acids (and sterols) have been all but ignored by biochemists; these polyannular molecules are considered laboratory products and they play no part in the biosynthesis of bile acids and sterols. The latter are synthesized by processes, still obscure, from simple molecules whose nature is yet a matter of conjecture; but these syntheses presumably follow lines quite different from those encountered in the structural analysis. This explains why the knowledge of the physicochemical and biochemical properties of this group of derivatives is almost entirely limited to its natural representatives, the conjugated bile acids and the free bile acids.

##### *A. Analytical methods; color reactions*

Some of the most characteristic chemical and physical properties of bile acids are involved in the methods developed for their analysis (43).

The oldest reaction for bile acids is the Pettenkofer reaction (142).

Bile acids, when heated with sugar in concentrated sulfuric acid, develop a beautiful purplish coloration. This reaction is by no means specific and even the numerous applications of modern colorimetric and spectroscopic methods, the latter also applied to the fluorescence of the Pettenkofer pigment, have never led to an accepted standard method (5, 34). The reaction was improved by Udransky (186), who supplanted saccharose by furfural, by Jolles (103), who recommended rhamnose or methylfurfural, and by Inouye and Ito (100), who used vanillin. The presence of other potential contributors to a positive Pettenkofer test, as terpenes etc., has been excluded by chemical isolation of the bile acids as salts, for instance as quinine salt (191), prior to the performance of the color test. In some instances, the Pettenkofer reaction has resulted in interesting biological findings, but its quantitative evaluation has never escaped controversial ground.

Enderlen (62) observed a brilliant blue color reaction when replacing the concentrated sulfuric acid by dilute, yet strong, sulfuric acid. This reaction has been rediscovered by Gregory and Pascoe (80), and studied by Reinhold and Wilson (152) and by Doubilet (58). It is specific for cholic acid and its conjugated derivatives; deoxycholic acid and other dihydroxycholanolic acids and their conjugation products fail to produce the blue substance with furfural and 70 per cent sulfuric acid except for faint traces which might still be due to contamination with cholic acid. Whereas this method promises to permit differentiation among the various bile acids, one has to turn to some other means for the determination of total bile acids. A general method, harboring this promise, has been proposed by Szillard (181); the insolubility of the ferric salts of bile acids and a quantitative determination of the iron by one or another titrimetric or colorimetric reaction forms a sound stoichiometric basis for quantitative estimation.

Liebermann's color reaction with acetic anhydride and concentrated sulfuric acid, used widely in the analysis of cholesterol, is also given by a number of bile acid derivatives, but can not be used for their estimation. (229, p. 277).

A very valuable aid in the analysis of conjugated bile acids is the determination of amino nitrogen by van Slyke's method, as introduced by Foster and Hooper (67) and Schmidt and Dart (168). Its results have to be interpreted with discrimination, since the possibility of the presence of non-conjugated bile acids in physiological and pathological specimens can not be dismissed (77, 169).

### *B. Physical properties*

An analytical method based on physical properties of the bile acids is the polarimetric determination of purified extracts (102). The value of

this procedure is impaired by the differences in specific rotation between the bile acids occurring in varying proportion in biological fluids where quantitative estimation is needed.

The high viscosity of blood serum in jaundice is due to the presence of bile acids, but the quantitative relations of this property are not sufficiently clarified as to permit interpretation in terms of bile acid content.

The immense surface activity of bile acids has been utilized for a long time in the so-called Hay test (92, 72). Flowers of sulfur is strewn on the surface of a fluid suspected to contain bile acids. The specific gravity and the size of the particles of the sulfur are such that they will float on water ( $s = 73$  dynes per square centimeter) and on liquids with a moderately reduced surface tension as long as the latter is above 52 to 56 dynes per square centimeter. If the surface forces drop below this value, they cannot hold the sulfur particles which, in consequence, sink to the bottom. The positive outcome of this test will establish a minimum content, in terms of one of the bile acids chosen as a standard, and subject to the nature and amount of other constituents of the sample, especially proteins. Quantitative surface tension measurements can be made with a variety of procedures such as capillary methods, stalagmometric methods, and especially by estimation of the static surface tension, using a ring, blade or stirrup (61, 78) in connection with a mechanical or electrical torsion balance. Influence of interfering substances, such as polyvalent electrolytes (phosphates) or colloids (proteins), can be eliminated by operation in an acid medium, and by comparison with a blank measured after absorption of the bile acid on charcoal (60). Comparisons of the surface activity of free and unconjugated bile acids will shed further light on the analytical possibilities of this property (59, 63).

### *C. Monomolecular layers*

The progress of our knowledge of monomolecular layers will eventually allow an interpretation of the data on surface forces in the light of molecular dimensions. Studies on monomolecular layers have already been undertaken by Adam, Rosenheim, and Danielli (1, 2, 51, 52) in the sterol series, and their results regarding the cross section of the sterol molecule are in perfect agreement with x-ray data. Views regarding the tilt of the molecules in such layers, the polarity of the functional groups, the position of double bonds, as revealed by their behavior when spread on permanganate solutions, are additional achievements of these studies.

### *D. X-ray analysis*

As mentioned in the introduction, x-ray spectroscopy and optical crystallography "pointed clearly to the fact that the older accepted formulas

(sc. for the sterols) could not be made to fit into the crystallographic cell" (13). In the case of ergosterol, for instance, the formula proposed by Wieland and Windaus requires cell dimensions of  $8.5 \times 7.0 \times 18.0$  A.U. or according to Wieland's modified formula (228)  $11.0 \times 7.5 \times 15.0$  A.U. Rosenheim and King's formula fits into a cell of  $7.5 \times 4.5 \times 20.0$  A.U., while Bernal observed  $7.2 \times 5.0 \times 20.0$  A.U. for ergosterol. Similar confirmations of Rosenheim and King's theory are obtained from the molecular dimensions of pregnandiol and ketohydroxyoestrin. Cholic acid does not give such clear evidence, as the packing is more complex owing to the various functional groups. However further study, especially of the underlying hydrocarbons, promises a general elucidation of these points and a correlation with other phenanthroid nuclei as in Jacobs' hydrocarbon  $C_{16}H_{14}$  (101) from strophanthidin (Bernal in reference 52). For x-ray measurements of cholic acids see p. 360.

The results of ultra-violet spectroscopy have been consulted with great advantage in the study of unsaturated sterol compounds, especially ergosterol and its irradiation products (T. M. Lowry in reference 52). An investigation by Menschick, Page, and Bossert (133) had a more direct bearing on the chemistry of bile acids. They concluded from the ultra-violet absorption curve of coprostenone (old nomenclature: "cholestenone") (XV), that it belongs to the type of  $\alpha, \beta$ -unsaturated and not to that of  $\beta, \gamma$ -unsaturated ketones. This fact could not easily be reconciled with the old sterol formula, but its implications regarding oxidative destruction of coprostenone dovetail with all other evidence in favor of the new formula.

The effects of irradiation on ergosterol are beyond the scope of this review. Its effects on the steric configuration of cholic acid described by Uraki (187) require further confirmation.

#### *E. Acid dissociation constants*

Qualitative observations on the relative acidity of bile acids have been utilized in the separation of lithocholic acid from the more acid deoxycholic acid (see p. 333). Phocaecholic acid, an  $\alpha$ -hydroxy acid, would be expected to be among the strongest acids in this group.

Quantitative studies on the acid dissociation of bile acids are scarce. Bondi, from conductivity measurements (19), estimated the dissociation constant of glycocholic acid at  $k = 1.32 \times 10^{-4}$  ( $pK = 3.87$ ). He gives  $1.38 \times 10^{-4}$ ,  $0.18 \times 10^{-4}$ , and  $2.2 \times 10^{-4}$  for the dissociation constants of lactic, acetic, and hippuric acids, respectively. From the ratio of the constants for hippuric acid and benzoic acid ( $2.2 \times 10^{-4} : 0.6 \times 10^{-4}$ ), he deduces that the dissociation constant of free cholic acid must be less than  $0.6 \times 10^{-4}$  ( $pK = 4.22$ ). Henriques (95) used the method of Michaelis and Mizutani (134), which consists in the colorimetric comparison of an



unknown solution with a known buffer solution of glycine plus acetic acid in alcohol. He obtained  $pK = 2.78$  and  $3.23$  for glycocholic and taurocholic acid (Bondi's  $k = 1.32 \times 10^{-4}$  corresponds to a  $pK$  of  $3.87$ ) and  $2.46$  and  $2.93$  for "glycocholeic and taurocholeic acid," expressions that should be replaced by glycocodeoxycholic and taurodeoxycholic acid. Josephson (104), working with electrometric methods on aqueous solutions of perhaps purer materials, obtained the following values for  $pK$  (those marked with an asterisk corrected for activity coefficients): cholic acid,  $5.19$ – $5.20^*$ ; deoxycholic acid,  $6.42$ – $6.45^*$ ; glycocholic acid,  $4.45$ – $4.51^*$ ; glycocodeoxycholic acid,  $3.81$ ; taurocholic acid,  $1.39$ ; taurodeoxycholic acid,  $1.74$ . The low values (strong acidity) of the tauro acids are in good agreement with the older values of  $pK = 1.17$  for taurocholic acid ( $4.40$  for glycocholic acid), given by H. Hammarsten (86). These important constants deserve further study, e.g., by the method of Kuhn and Wassermann (110) for the determination of dissociation constants of acids scantily soluble in water.

#### VI. CHOLEIC ACIDS

We have reserved until here the discussion of one group of bile acids, the choleic acids. The term "Choleinsaeure," coined by Demarçay (45), had covered an ever-changing complex of bile educts (183) until it was fixed by Lachinov (114) upon a bile acid which accompanied cholic acid in varying proportion. This deconjugated, i.e., non-nitrogenous, bile acid showed certain similarities with Mylius' deoxycholic acid (135) and could be converted into deoxycholic acid containing "crystal-acetic acid" by treatment with glacial acetic acid. The question as to the identity or diversity of these two substances remained undecided, while the terms choleic and deoxycholic acid were used indiscriminately, e.g., by Wahlgren (194) and Gullbring (84), who isolated glycocholeic acid (= glycocodeoxycholic acid) and taurocholeic acid (= taurodeoxycholic acid) in Hammarsten's laboratory.

All incongruities were clarified at once through the ingenious conception of the "choleic acid principle" by Wieland and Sorge (200). In an attempt to study the effect of high vacuum distillation on choleic acid, a method which had proven useful in the case of cholic acid, these authors observed the appearance of palmitic, stearic, and oleic acids in the head fraction of the distillate to an extent of about 8 per cent of the choleic acid employed. The remainder proved identical with deoxycholic acid. Prompted by Werner's achievements in complex compound chemistry, they realized that they were dealing with a complex or coördination compound consisting of eight molecules of deoxycholic acid plus one molecule of fatty acid. This molecular compound had a higher melting point ( $186^\circ C.$ ) than either

of its constituents, deoxycholic acid melting at 172°C. It differed from it in its solubilities, but since the fatty acid could be replaced by an excess of acetic acid, acetone, alcohol, or ether, it can easily be converted into deoxycholic acid, which is known to crystallize with one molecule of one of the solvents mentioned for each molecule of bile acid. The complex "choleic acid" can be resynthesized from fatty acid and deoxycholic acid.

Further experiments showed that other fatty acids, as butyric acid, and also aromatic substances, like phenol and naphthalene, were able to enter coördination compounds with deoxycholic acid in varying but characteristic molecular proportions. The term "choleic acid" was now widened to comprise these addition compounds, and the choleic acid containing stearic, palmitic, or oleic acid, and isolated from animal bile, was designated "natural choleic acid."

To the formation of these compounds was ascribed the important physiological function of bile acids to render all sorts of water-insoluble substances water-soluble. Since the alkali salt, e.g., of naphthalenecholeic acid or camphorcholeic acid, is water-soluble, naphthalene, camphor, etc., can thus be held in aqueous solution. The significance of this principle for the theory of intestinal resorption is obvious. While the physical process of emulsification can not be ignored, any explanation of lipid resorption and transport would be incomplete without the choleic acid principle.

This peculiarity of deoxycholic acid is shared only to a limited extent by other bile acids. Cholic acid combines with various alcohols and mercaptans according to Mylius (135a); dehydrocholic acid crystallizes with benzene in the molecular proportion 2:1 (37). Various acetyl and halogen derivatives of hydroxy-, hydroxyketo-, and keto-cholanic acids crystallize with acetic acid, acetone, ethyl acetate, alcohol, ether, etc. But all these acids combine with the solvent in the molecular proportion 1:1; thus they do not distinguish themselves from many other known organic compounds.

The existence of unusual coördinative valence in this group is responsible for many double compounds such as "Weyland's acid," which was shown by Wieland and Kishi (234) to consist of one molecule each of anthropodeoxycholic (= chenodeoxycholic) acid and 3-hydroxy-12-ketocholanic acid. Likewise, in the sterol series,  $\gamma$ -cholestanol was recognized as a double compound of cholestanol with epicoprosterol (250). Page and Mueller (141) demonstrated the presence of a double compound of cholesterol and dihydrocholesterol in human brain. The combination of various sterols with digitonin and other saponins, discovered by Windaus (254), should be mentioned in this connection.

Returning to deoxycholic acid itself, one has to record that a great many careful observers before Wieland and Sorge had noted the stubborn tenacity by which fatty acids are retained whenever the purification of bile

acids had been attempted (e.g., 112, 113). This was perhaps due to the presence of deoxycholic acid, which in turn is not easily separated from cholic acid.

Boedecker and Volk (17) discovered that this property of deoxycholic acid was duplicated in the behavior of apocholic acid (but not of the isomeric dihydroxycholenic acid). This fact hints at the specific enhancing influence of the hydroxyl groups in positions  $C_3$  and  $C_{12}$  on coördinative valence. If we compare deoxycholic acid and apocholic acid on one side, with cholic acid, lithocholic acid, and cholanic acid on the other side, all of which have no pronounced coördinative power, we are led to the following assumption. The residual valences in the second group of acids are scattered in various directions from the main longitudinal axis of the molecule which stretches from the hydroxyl on  $C_3$  in lithocholic acid to the distal common carboxyl group in position 24 at the end of the side chain. In deoxycholic and apocholic acid, the second hydroxyl ( $C_{12}$ ) seems to direct these residual affinities to the opposite front of the molecule and we incline to the view that the unbroken chain of aliphatic carbons  $C_3$   $C_4$   $C_5$   $C_7$   $C_8$   $C_{14}$   $C_{15}$   $C_{16}$   $C_{17}$   $C_{20}$   $C_{22}$   $C_{23}$   $C_{24}$  is endowed with this particular clinging power. (This chain would be interrupted by a double bond between  $C_7$  and  $C_8$  in the dihydroxycholenic acid isomeric with apocholic acid.) The question why cheno- or hyo-deoxycholic acids do not exhibit similar affinity on the other side of the molecule (chain  $C_3$   $C_2$   $C_1$   $C_{10}$   $C_9$   $C_{11}$   $C_{12}$   $C_{13}$   $C_{17}$   $C_{20}$   $C_{22}$   $C_{23}$   $C_{24}$ ) finds its answer by inspection of the spatial model. This front of the molecule is interrupted by the methyl groups  $C_{18}$  on  $C_{10}$ ,  $C_{19}$  on  $C_{13}$ , and  $C_{21}$  on  $C_{20}$ . This hypothesis might be tested in the future by methods of "molecular anatomy." A start in this direction was made by Herzog, Kratky, and Kuriyama (96), who studied the x-ray structure of stearic acid-choleic acid, with the result that the crystal cell was found only to occupy one-half of the minimum molecular volume as computed from molecular weight and density of a stearic acid-octocholeic acid. Perhaps the eight molecules of bile acid are arranged in two tetrads around the upper and lower half of the stearic acid chain, allowing two "identities" or pseudo-identities within one molecule of the coördination compound. Further studies on the crystal structure of choleic acids, carried out by Go and Kratky (76), indicate that crystals of choleic acids of coördination number 2 and 6 likewise contain cells, comprising four molecules of deoxycholic acid, thus corresponding to two molecules, and  $\frac{2}{3}$  of a molecule, respectively, of the choleic acid.

Which substances function as partners in choleic acids? A great many compounds have been studied, especially by Rheinboldt and his coworkers (153) and by Sobotka, Goldberg, and Kahn (175, 176, 177), who sum-

marize them under the term of "acholic constituent" by analogy with the "aglucone" of glucosides.

#### A. *Coördination number*

While Wieland and Sorge originally assumed that the number of deoxycholic acid molecules combining with the monocarboxylic acids of the aliphatic series was in the simple proportion of one molecule of bile acid for each  $-\text{CH}_2-\text{CH}_2-$ , systematic research by Rheinboldt showed that only those coördination numbers can be verified in choleic acids which have been recognized as the general architectural principle in inorganic coördination compounds since Werner. These numbers must allow for a symmetrical arrangement around the pivotal atom, ion, or group. The numbers commonly encountered are 4, 6, 8, occasionally 2, and 3, while the numbers 5 and 7 would not permit symmetrical constellation. Potential higher coördination numbers are 12 and 20, according to Hüttig (98), but no evidence has been adduced for the existence of such huge symmetrical molecular aggregates. Figure 1 gives the coördination number of homologous aliphatic monocarboxylic and dicarboxylic acids. Physical properties in homologous series, such as melting points, solubility, acid dissociation constant, surface activity, etc., change regularly and gradually with the molecular size, i.e., with the number of carbon atoms, or follow an oscillatory course between the homologs with an even and those with an odd number of carbon atoms. In contradistinction, coördinative valence in an homologous series is not a gradual, but a periodical function of molecular size and to such a factor one might ascribe peculiarities and specificities in the biochemical behavior of homologous members.

Rheinboldt found that an alcohol with  $n$  carbon atoms has the same coördination number as the acid with  $(n + 1)$  carbon atoms. In alkyl esters of aliphatic acids, the coördination number peculiar to the longer of the two chains prevails. Halogenated aliphatic acids (and in some instances unsaturated acids) behave as a rule like the saturated acids of equal carbon number. The influence of branched chains upon coördinative valency was studied by Sobotka and Goldberg, who found the coördination number reduced to that of the longest straight chain contained in the molecule. The coördination number 3 was observed in crotonic acid, in some enol derivatives, and in aromatic substances such as phenol. Whether this is a peculiarity of unsaturated and aromatic compounds must be decided by a survey of a greater number of choleic acids. Chargaff and Abel (278) likewise found decreased coördination numbers for  $\alpha$ -alkylated fatty acids; they obtained no choleic acids at all with halogenated fatty acids.

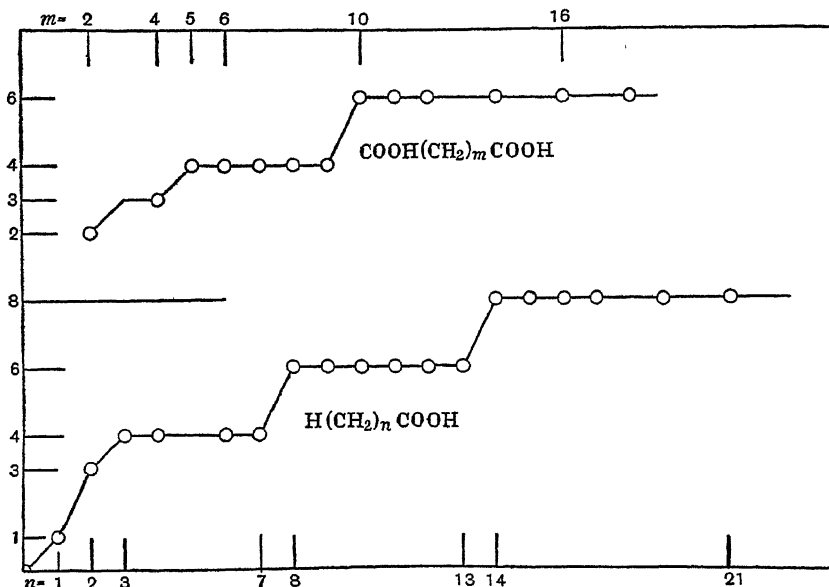


FIG. 1. RELATION BETWEEN LENGTH OF ALIPHATIC CHAIN AND COÖRDINATION NUMBER

Sobotka and Goldberg (175)

### B. Coördination compounds and optical activity

While the determination of coördinative valence towards bile acids offers a quantitative approach to several problems, there are additional features of even greater significance to be taken into account. Choleic acids vary not only as to coördination number but also as to their stability. This is a function of volatility and solubility of the acholic component, and of the degree of "saturation" of the coördinative valence. One acholic constituent might dislodge another owing to its greater coördinative affinity, a reaction that has been utilized by Wieland and by Rheinboldt in the analysis of aliphatic acid-choleic acids which dissociate in xylene to insoluble xylenecholeic acid and free aliphatic acid. Differences in physical properties enabled Sobotka and Goldberg (176) to isolate choleic acids of levorotatory camphor, phenylethylethanol, dipentene, and methyl-ethylacetic acid from solutions of the respective racemic acholic components and deoxycholeic acid. Owing to the optical activity of the bile acid, *l*-camphorcholeic acid and *d*-camphorcholeic acid are not mirror images of each other, but differ in their physical properties and can thus be separated by crystallization. This phenomenon offers a possibility of resolving racemic hydrocarbons and other substances, devoid of functional groups that would allow resolution by the customary Pasteurian methods.

### *C. Multiple coördination numbers*

These have been encountered in several acholic substances; in other words, some compounds form more than one choleic acid. Rheinboldt observed the existence of camphormonocholeic and camphordicholeic acids (154). Depending on the method of synthesis, two choleic acids each were obtained by Sobotka and Goldberg from butyric acid and from its ethyl ester. The butyric acid-choleic acid with coördination number 2 was prepared by direct synthesis, i.e., by dissolution of solvent-free deoxycholic acid in butyric acid. If, on the other hand, butyric acid and deoxycholic acid are combined in alcoholic solution, without an excess of butyric acid, then four molecules of deoxycholic acid will combine with a single molecule of butyric acid. Thus, the initial concentration of the two compounds influences the coördinative ratio.

The presence of alcohol or some such solvent seems to compete with the acholic constituent for the deoxycholic acid; therefore, many substances can not be induced to form choleic acids in alcoholic solution, e.g., glutaric acid and various branched-chain aliphatic acids. The direct method of synthesis is indispensable in these cases. In instances of multiple coördinative valence, the direct method will yield the compound with lower coördination number.

A third method of synthesis consists in the addition of the barely water-soluble acholic component, e.g., ethyl butyrate, to an aqueous solution of sodium deoxycholate. The lipid phase solidifies gradually and the resultant crystalline layer is ethyl butyrate-dicholeic acid. This last method also illustrates how the addition of a non-electrolyte phase to a bile salt solution influences the ionic concentrations of the aqueous phase through hydrolytic absorption. The sodium ion is left in the aqueous phase and the pH of this phase is raised.

### *D. Tautomerism and coördination*

The study of ethyl acetoacetate choleic acid and the choleic acids of related tautomeric diketones yielded results of twofold interest (177). While ethyl acetoacetate in liquid form or in solution contains between 8 and 13 per cent of enol, it is 100 per cent enolized in its choleic acid. The corresponding figures for acetylacetone are 70 per cent enol in solution and 100 per cent in combination with deoxycholic acid. Through the crystallization of the choleic acid the tautomeric equilibrium is disturbed and the entire amount is gradually enolized. When the crystals of these choleic acids are dissolved, they undergo molecular dissociation and the liberated enol is ketonized at a measurable rate until the tautomeric equilibrium is established. Very quick operation however shows that maximal enol percentages approximating 100 per cent are reached only 30 seconds to a

minute after dissolution, which indicates that molecular dissociation proceeds with a high but measurable speed.

The assumption that choleic acids when dissolved are subject to complete or at least partial molecular dissociation is further supported by molecular weight determinations. Such determinations were carried out by Sobotka and Goldberg in Signer's apparatus (174), which utilizes the principle of isothermic distillation of Barger, on sebatic acid-choleic acid. The apparent molecular weight in methyl alcohol indicates practically complete dissociation into sebatic acid and eight molecules of deoxycholic acid.

Both experiments, the titration of free enol in the case of acetoacetic ethyl ester, and the molecular weight determination in the preceding example, were performed in alcoholic solution, thus in a solvent which has an affinity of its own towards bile acid. Different conditions may exist in aqueous solution, where interrelation of molecular dissociation and electrolytic dissociation of the bile acid and also of the acholic component, if an acid, have to be considered (cf. p. 363, synthesis from aqueous solution). This subject is of great importance for the application of Wieland's choleic acid principle to alimentary adsorption.

#### *E. Resorption in the intestinal tract*

One has to keep in mind that animal bile contains conjugated bile acids. No choleic acids with conjugated bile acids have been unequivocally isolated (170), but the conjugated acids share the ability of the free bile acids of keeping water-insoluble substances in solution. The capacity of free bile acids to keep fatty acids and other acholic components in aqueous solution, whether due to actual coördination compounds or to surface forces, is confined to the alkaline side of the neutral point (189). Conjugated bile acids, owing to their greater acidity, retain this dissolving power even in slightly acid reaction down to pH 6.

Another angle from which this problem can be approached is the study of diffusion in mixtures containing bile acids (188, 73). The present state of knowledge is still limited, as the influence of pH, the nature of the membrane, and several other factors have to be considered individually. One may surmise that combination with bile acids will increase the diffusibility of substances less diffusible than the bile acids themselves, while it will cause the retention of substances more diffusible. Furthermore, the influence of the bile acids will be enhanced towards more alkaline solutions. These conditions may be complicated in the animal body (a) by enzymatic reactions disturbing the equilibrium on either side of the membrane, e.g., hypothetical resynthesis of neutral fat within the duodenal villi, (b) by the peculiar physical properties of such substances as lecithin, which exert an influence of their own on solubility in mixtures and may supplement the

effect of the bile acids (73). The observation that bile acids may promote the diffusion of multiple amounts of lipoids through the intestinal mucosa may be explained thus: the bile acid complex is absorbed on the membrane, and the acholic component is removed by diffusion, while the bile acid returns into the solution (189).

The resorption of drugs, vitamins, and hormones is also linked up with the circulation of the bile. Wieland suggested that the resorption of alkaloids such as strychnine in the alkaline intestinal tract is based on the choleic acid principle. (Cf. e.g., Kolda (107) for alkaloids; Mendel and Daniels (132) and Stockholm and Schmidt (180) for fat-soluble dyes; Crowe (42) for hexamethylenetetramine). This idea has been utilized in the simultaneous administration of synthalin with sodium dehydrocholate. The failure of the peroral administration of insulin-bile salt mixtures to produce hypoglycemia (179, 74) is not surprising when one realizes the chemical nature of this hormone. On the other hand, the female sex hormone, owing to its lipid character, has a definite affinity towards bile: it has been detected in the gall bladder of an old man (70, 81) and its ability to form a choleic acid was put to preparative use by Wieland, Straub, and Dorfmueller (239). The relationship between vitamin A and bile acids postulated by Shimizu and Hatakeyama (172) must be examined in the light of the modern knowledge of this vitamin. Seyderhelm (171) demonstrated the necessity of peroral administration of bile acids for the resorption of the fat-soluble vitamin D to counteract the typical anemia developing in dogs whose bile flow is diverted from the intestine through a bile fistula.

The enolizing influence of bile acids, mentioned above, has been adduced to explain the difference between the direct and indirect diazo reaction of bilirubin according to Hymans van den Bergh. Bilirubin in the serum of jaundiced subjects seems to occur in two forms. One form, suspected to be the enol, couples readily with diazotized sulfanilic acid, while the other one reacts only after it has been extracted with alcohol. A study of the incidental pathological conditions suggests that the simultaneous presence of bile acids is prerequisite for the "direct" reaction, while the less frequent sera with "indirect" reaction only contain the bilirubin in the keto form and require the alcohol treatment for enolization (69, 177). The enolizing influence of bile acids may play a part in the beta-oxidation of fatty acids.

The influence of the bile acids on enzymatic reactions is especially noticeable with the lipolytic enzymes. Lipase, the enzyme splitting triglycerides, is enhanced in its action by bile and bile acids, while esterase, the intracellular enzyme of the liver, which is most active towards simple aliphatic esters, is inhibited by bile salts. This antagonism may find its explanation in the surface activity of the bile acids, but since hydrolysis might be



preceded by enolization of the  $\text{—CH}_2\text{—C}\begin{smallmatrix} \text{O} \\ \parallel \\ \text{OEt} \end{smallmatrix}$  group to  $\text{—CH=}\begin{smallmatrix} \text{OH} \\ \diagup \\ \text{OEt} \end{smallmatrix}$ , we

can conceive of more specific reasons for these phenomena.

#### VII. CONCENTRATION AND pH OF BILE

Bile as excreted from the liver displays an alkaline reaction, varying from 7.8 to 8.6 in various species (see references 137, 139, 146). It contains on the average 1 to 4 per cent of solid constituents; these are bile acids in the form of their sodium salts (in some marine animals also a high amount of potassium salt), minor amounts of lecithin and cholesterol, inorganic constituents of the same type as in blood serum, with sodium chloride prevailing, and fluctuating amounts of mucin, in part contributed by the lining of the bile ducts. This liver bile is stored in many species in the gall bladder, where it undergoes concentration at a remarkable rate (162). Gall-bladder bile might contain as much as 20 per cent of total solids as a result of a tenfold concentration within a few hours. It seems paradoxical, then, that both liver bile and gall-bladder bile should be isotonic with blood serum and the other body fluids. But this fact has been ascertained by Brand (27) through freezing-point determinations. Brand revealed that the concentration of the inorganic constituents is gradually reduced through diffusion from 0.8 per cent to 0.2 per cent, while the organic components increase. These conditions must be governed by Donnan's law, as we are dealing with a mixture of sodium chloride and sodium bile salts within the bladder and a solution of sodium chloride in the outside fluid, the blood of the capillary system of the bladder wall that takes up the fluid. Although the constant flow of the blood and of the bile prevents the establishing of a true equilibrium, the bladder bile turns more acid and, in concentrated state, reaches pH values below 6 (151). The influence of this shift of sodium and hydrogen ions upon the capacity of the bile to keep lipoids, especially cholesterol, in the dissolved state and upon its diffusibility is not yet understood. We know only that those human gall bladders in which stones are found at autopsy regularly show a ratio bile acids: cholesterol of less than 8:1 by weight, while stone-free bile contains a higher relative amount of bile acid, regardless of the absolute amount of solid ingredients in either group (136).

#### VIII. PHYSIOLOGICAL AND PHARMACOLOGICAL OBSERVATIONS

Bile participates in the detoxifying function of the liver, in the case of both organic and inorganic poisons. There are numerous references to the

presence of heavy metals in the bile in poisoning (lead (7), manganese (148), cresol as glucuronic derivative (16)). The affinity of the bile for foreign organic substances is best exemplified by the accumulation of tetraiodophenolphthalein and other radiopaque chemicals, used in the visualization of this organ, and of organic mercury compounds and other disinfectants which were introduced into the body with the aim of sterilizing the infected gall bladder (typhoid carriers, etc. (138)).

The physiological and pharmacological literature abounds with studies on bile fistulas. There is hardly a group of substances that has not been tried for its cholagogue, choleric, or chologenic value. Choleric action designates the stimulation of the liver to excrete more bile, while cholagogue action refers to compounds which stimulate the emptying of the gall bladder. Of greater interest to the chemist is the study of chologenic substances, namely, substances that can be identified as actual starting material from which bile acids are synthesized. Such synthesis has not been demonstrated up to the present time, except for the observation that the output of taurocholic acid depends on the supply of cystine sulfur (11, 68). Despite the absence of conclusive physiological evidence, all that has been disclosed in regard to the constitution of sterols and bile acids, especially the new data on "oxycholesterol" (160), will induce the chemist to suspect that bile acids are formed in the liver by oxidative processes from a compound closely related to cholesterol.

Bile under physiological conditions is confined to the well-known cycle liver-gall bladder-duodenum-lymph spaces-portal system-liver. Under pathological conditions, the bile acids as well as other bile constituents, will trespass into blood serum and urine. The presence of bile acid in urine has been demonstrated by many methods. Its occurrence in the blood is inferred by its appearance in the kidney excretion, but its direct quantitative estimation in the blood depends on the establishment of more exact analytical methods (see pp. 354 to 356).

The toxic action of bile acids on the animal organism has been studied by the pharmacologist Hermann Wieland (242) by intravenous injection into animals or surviving preparations of heart and muscles. Especially does its effect on the heart resemble that of other surface-active substances like the saponins. Its lytic effect on single cells, as the red blood corpuscles or the pneumococcus, while of practical importance, is non-specific and quantitatively inferior to the lytic action of fatty acids. Deoxycholic acid produces greater effects than cholic acid in most of the biological phenomena enumerated, but we hesitate to ascribe this solely to its greater coördinative affinity. On the other hand, deoxycholic acid is more easily detoxified by the serum proteins than cholic acid, so that their final action, e.g., on the heart after intravenous injection, is about equal, according

to Hermann Wieland. The part played by the white blood corpuscles in the detoxification of bile acids has to be further investigated (41). The use of hemolysis and other biological phenomena in the analysis of bile acids, as suggested by some authors, can not be recommended.

The final fate of the bile acids in the animal organism is even more obscure than their origin. The tetracyclic carbon skeleton of the sterols, the bile acids, and the sex hormones has earned its wide-spread occurrence in nature by its great stability. It is hardly attacked by microorganisms and will persist in the carcass as adipocere long after carbohydrates, lipoids, and proteins have fallen a prey to the various saprophytes. No fragments that would offer a clue to the manner of its catabolism have been found in the living or dead animal body nor in its excretion products.

#### IX. CONCLUSION

We hope to have demonstrated in the foregoing pages (a) that we now possess a formula for the bile acids and related substances which agrees with the known chemical and physical facts and harmonizes with our general conceptions of the structure of natural substances, (b) that the possibility of syntheses in this group has been brought within our grasp, (c) that natural synthesis and catabolism of bile acids is still terra incognita, (d) that the establishment of the constitution opens the way for intelligent interpretation and research of the physicochemical properties of bile acids, and (e) that such knowledge will cause the numerous physiological, pharmacological, and pathological facts and observations, accumulated in the past, to be understood and correlated in the future.

#### REFERENCES

- (1) ADAM, N. K., AND ROSENHEIM, O.: *Proc. Roy. Soc. London* **126A**, 25 (1929).
- (2) ADAM, N. K., DANIELLI, J. F., HASLEWOOD, G. A. D., AND MARRIAN, G. F.: *Biochem. J.* **26**, 1233 (1933).
- (3) ADAM, N. K., ET AL.: *Nature* **132**, 205 (1933). Nomenclature of the oestrin group.
- (4) ADICKES, F.: *Ber.* **53**, 211 (1925).
- (5) ALDRICH, M., AND BLEDSOE, M. S.: *J. Biol. Chem.* **77**, 519 (1923).
- (6) ANDERSON, R. J.: *J. Am. Chem. Soc.* **46**, 1717 (1924); **48**, 2972 (1926).
- (7) ANNUSCHAT, A.: *Arch. exptl. Path. Pharmacol.* **7**, 45 (1877).
- (8) ASCHAN, O.: *Ber.* **45**, 1604 (1912).
- (9) BAEYER, A. VON, AND VILLIGER, V.: *Ber.* **32**, 3625 (1899).
- (10) BERGMANN, E., AND HILLEMANN, H.: *Ber.* **66**, 1302 (1933).
- (11) BERGMANN, G. VON: *Hofm. Beitr.* **4**, 192 (1903).
- (12) BERGMANN, W.: *J. Biol. Chem.* **104**, 317, 553 (1934).
- (13) BERNAL, J. D.: *Nature* **129**, 277 (1932); *Chemistry & Industry* **51**, 466 (1932).
- (14) BERNAL, J. D., AND CROWFOOT, D.: *Chemistry & Industry* **52**, 729 (1933).
- (15) BLANC, H. G.: *Compt. rend.* **144**, 1356 (1907).
- (16) BLUMENTHAL, F.: *Biochem. Z.* **1**, 135 (1906).
- (17) BOEDECKER, F., AND VOLK, H.: *Ber.* **53**, 1852 (1920); **54**, 2302, 2489 (1921).

- (18) BONDI, S., AND MÜLLER, E.: *Z. physiol. Chem.* **47**, 499 (1906).  
(19) BONDI, S.: *Z. physiol. Chem.* **53**, 8 (1907).  
(20) BONDZYSKI, ST., AND HUMNICKY, V.: *Z. physiol. Chem.* **22**, 396 (1897).  
(21) BORSCHÉ, W., AND ROSENKRANZ, E.: *Ber.* **52**, 342 (1919). Untersuchungen über der Konstitution der Gallensäuren. Paper No. 1.  
(22) BORSCHÉ, W., AND FRANK, R.: *Ber.* **59**, 1748 (1926). Paper No. 9.  
(23) BORSCHÉ, W., AND FRANK, R.: *Ber.* **60**, 723 (1927). Paper No. 10.  
(24) BORSCHÉ, W., AND FRANK, R.: *Ber.* **60**, 1216 (1927). Paper No. 11.  
(25) BORSCHÉ, W., AND FESKE, E.: *Z. physiol. Chem.* **176**, 109 (1928). Paper No. 13.  
(26) BORSCHÉ, W., AND TODD, A. R.: *Z. physiol. Chem.* **197**, 173-90 (1931). Paper No. 15.  
(27) BRAND, J.: *Pflüger's Arch.* **90**, 491 (1902).  
(28) BRAUN, J. VON, AND IRMISCH, G.: *Ber.* **65**, 883 (1932).  
(29) BREDT, J.: *Ann.* **437**, 1 (1923).  
(30) BUTENANDT, A.: *Ber.* **63**, 659 (1930); **64**, 2529 (1931).  
(31) BUTENANDT, A., WEIDLICH, H. A., AND THOMPSON, H.: *Ber.* **66**, 601 (1933).  
(32) CALLOW, R. K., AND ROSENHEIM, O.: *Chemistry & Industry* **52**, 272 (1933); *J. Chem. Soc.* **1933**, 387.  
(33) CERECEDO, L. R.: Thesis, Munich, 1921.  
(34) CHARLET, M.: *Biochem. Z.* **210**, 42 (1929).  
(35) CLEMMENSEN, E.: *Ber.* **46**, 1839 (1913); **47**, 51, 681 (1914).  
(36) CLEVE, P. T.: *Bull. soc. chim.* **35**, 373, 429 (1881); **38**, 131 (1882).  
(37) COHN, LASSAR: *Ber.* **25**, 803 (1892).  
(38) COOK, J. W., DODDS, E. C., AND HEWETT, C. L.: *Nature* **131**, 56 (1933).  
COOK, J. W., DODDS, E. C., HEWETT, C. L., AND LAWSON, W.: *Proc. Roy. Soc. London* **114B**, 272 (1934).  
(39) COOK, J. W., AND DODDS, E. C.: *Nature* **131**, 205 (1933).  
COOK, J. W. (in discussion on tumors): *Proc. Roy. Soc. London* **113B**, 273 (1933).  
(40) COOK, J. W., AND HEWETT, C. L.: *Chemistry & Industry* **52**, 451 (1933); *J. Chem. Soc.* **1933**, 1098.  
(41) CROFTAN, A. C.: *Am. J. Med. Sci.* **123**, 150 (1902); *Pflüger's Arch.* **90**, 635 (1902).  
(42) CROWE, J. S.: *J. Biol. Chem.* **4**, xxxv (1908).  
(43) CUNY, L.: *Le dosage des sels biliaires dans la bile, etc.*, Paris, 1930.  
(44) DANÉ, E., AND WIELAND, H.: *Z. physiol. Chem.* **194**, 119 (1931). Wieland, Paper No. 32.  
(45) DEMARÇAY, H.: *Ann.* **27**, 270 (1833).  
(46) DIELS, O., AND ABDERHALDEN, E.: *Ber.* **36**, 3177 (1903); **37**, 3092 (1904).  
(47) DIELS, O.: *Ber.* **58**, 1231 (1925); **60**, 140 (1927).  
DIELS, O., GÄDKE, W., AND KÖRDING, P.: *Ann.* **459**, 1 (1927).  
(48) DIELS, O., AND KARSTENS, A.: *Ann.* **478**, 136 (1930).  
(49) DIELS, O.: *Ber.* **66**, 487 (1933).  
(50) DIELS, O.: *Ber.* **66**, 1122 (1933).  
(51) Discussion on the Chemistry of Sterols: *Chemistry & Industry* **52**, 10 (1933).  
(52) Discussion on the Chemistry of Oestrin: *Chemistry & Industry* **52**, 268, 287 (1933).  
(54) DORÉE, CH., AND GARDNER, J. A.: *J. Chem. Soc.* **93**, 1328 (1908).  
(55) DORÉE, CH., AND GARDNER, J. A.: *J. Chem. Soc.* **93**, 1625 (1908).  
(56) DORÉE, CH.: *J. Chem. Soc.* **95**, 638 (1909).

- (57) DORÉE, CH., AND ORANGE, L.: J. Chem. Soc. **109**, 46 (1916).  
(58) DOUBILET, H.: Private communication.  
(59) DOUMER, E., AND DOUMER, E.: Compt. rend. soc. biol. **84**, 393, 683 (1921).  
(60) DUCCO, C. L., AND PANZA, P. T.: Semana méd. **37**, I, 1193 (1930).  
(61) DUÑOÛY, P. LECOMTE: Surface Equilibria of Biological and Organic Colloids. New York (1926).  
(62) ENDERLEN, C.: Ann. **75**, 174 (1850).  
(63) ETTISCH, C., AND KOGANEI, R.: Biochem. Z. **193**, 390 (1928).  
(64) FERNHOLZ, E.: Ann. **507**, 128 (1933); **508**, 215 (1934).  
(65) FISCHER, H.: Z. physiol. Chem. **73**, 204 (1911).  
(66) FLINT, AU., JR.: Am. J. Med. Sci. **44**, 305 (1862).  
(67) FOSTER, M. G., AND HOOPER, C. W.: J. Biol. Chem. **38**, 355 (1919).  
(68) FOSTER, M. G., HOOPER, C. W., AND WHIPPLE, A.: J. Biol. Chem. **38**, 379 (1919).  
(69) FOWWEATHER, F. S.: Biochem. J. **26**, 165 (1932).  
(70) FRANK, R. T., AND GOLDBERGER, M. A.: Private communication.  
(71) FRANK, R. T., AND SOBOTKA, H.: Proc. Soc. Exptl. Biol. Med. **29**, 1026 (1932).  
(72) FRENKEL, H., AND CLUZET, J.: Compt. rend. soc. biol. **52**, 1086 (1900).  
(73) FÜRTH, O., AND SCHOLL, R.: Biochem. Z. **222**, 430 (1930).  
FÜRTH, O., AND MINIBECK, K.: Biochem. Z. **237**, 139 (1931).  
(74) GAIS, E.: Proc. Soc. Exptl. Biol. Med. **29**, 216 (1932).  
(75) GIRARD, A.: Bull. soc. chim. biol. **15**, 562 (1933).  
(76) GO, Y., AND KRATKY, O.: Z. physik. Chem. **26B**, 439 (1934).  
(77) GOFF, M., HRDINA, L., AND ANDREWS, E.: Proc. Soc. Exptl. Biol. Med. **29**, 549 (1932).  
(78) GOIFFON, R.: Bull. soc. chim. biol. **8**, 165 (1926).  
(79) GRASSMANN, W., AND BASU, K. P.: Z. physiol. Chem. **198**, 193 (1931).  
(80) GREGORY, R., AND PASCOE, P. A.: J. Biol. Chem. **83**, 35 (1929).  
(81) GSELL-BUSSE, M. A.: Arch. exptl. Path. Pharmacol. **138**, 328 (1929).  
(82) GUITERAS, A.: Ann. **494**, 116 (1932).  
(83) GUITERAS, A.: Z. physiol. Chem. **214**, 89 (1933).  
(84) GULLBRING, A.: Z. physiol. Chem. **45**, 448 (1905).  
(85) GUNDELACH, C., AND STRECKER, A.: Ann. **62**, 205 (1847).  
(86) HAMMARSTEN, H.: Biochem. Z. **147**, 481 (1924).  
(87) HAMMARSTEN, O.: Ber. **14**, 71 (1881).  
(87a) HAMMARSTEN, O.: Z. physiol. Chem. **24**, 322 (1898).  
(88) HAMMARSTEN, O.: Z. physiol. Chem. **61**, 454 (1909).  
(89) HARPER, S. H., KON, G. A. R., AND RUZICKA, F. C. J.: J. Chem. Soc. **1934**, 124.  
(90) HARRIES, C.: Ber. **45**, 936, 943 (1912).  
(91) HAWORTH, R. D., ET AL.: J. Chem. Soc. **1932**, 1125, 1784, 2248; **1933**, 1012.  
(92) HAY, M.: in L. Landois and Stirling's Textbook of Human Physiology, 2nd American edition, p. 294. Philadelphia (1886).  
(93) HEILBRON, I. M., AND SEXTON, W. A.: J. Chem. Soc. **1924**, 2255.  
(94) HEILBRON, I. M., SIMPSON, J. C. E., AND SPRING, F. S.: J. Chem. Soc. **1933**, 626.  
(95) HENRIQUES, O. M.: Acta Path. Microbiol. Scand., Supplement **3**, 141 (1930).  
(96) HERZOG, R. H., KRATKY, O., AND KURIYAMA, S.: Naturwissenschaften **19**, 524 (1931).  
(97) HUECKEL, W., DANNEEL, R., SCHWARTZ, A., AND GEHRKE, A.: Ann. **474**, 121 (1929).

- (98) HÜTTIG, G. F.: *Z. anorg. Chem.* **114**, 24 (1920).  
(99) INHOFFEN, H. H.: *Ann.* **497**, 130 (1932).  
(100) INOUE, K., AND ITO, H.: *Z. physiol. Chem.* **57**, 313 (1908).  
(101) JACOBS, W. A., AND FLECK, E. E.: *J. Biol. Chem.* **97**, 57 (1932).  
(102) JENKE, M.: *Arch. exptl. Path. Pharmacol.* **130**, 280 (1928).  
(103) JOLLES, A.: *Z. physiol. Chem.* **57**, 30 (1908).  
(104) JOSEPHSON, B.: *Biochem. J.* **263**, 428 (1933).  
(105) KENNAWAY, E. L., AND COOK, J. W.: *Chemistry & Industry* **51**, 521 (1932).  
(106) KISHNER, N.: *J. Russ. Phys. Chem. Soc.* **44**, 861 (1912).  
(107) KOLDA, J.: *Compt. rend. soc. biol.* **94**, 216 (1928).  
(108) KON, G. A. R.: *J. Chem. Soc.* **1933**, 1081.  
(109) KOTAKE, M.: *Ann.* **465**, 1 (1928).  
(110) KUHN, R., AND WASSERMANN, A.: *Helv. Chim. Acta* **11**, 31 (1928).  
(111) LACHINOV (LATSCHINOFF), P. A.: Quoted from *Ber.* **12**, 1518 (1879).  
(112) LACHINOV, P. A.: *Ber.* **13**, 1911 (1880).  
(113) LACHINOV, P. A.: *Ber.* **15**, 713 (1882).  
(114) LACHINOV, P. A.: *Ber.* **18**, 3039 (1885).  
(115) LACHINOV, P. A.: *Ber.* **19**, 521 (1886).  
(116) LACHINOV, P. A.: *Ber.* **20**, 1053 (1887).  
(117) LANDSTEINER, K.: *Z. physiol. Chem.* **19**, 285 (1894).  
(118) LAPWORTH, A., AND CHAPMAN, E. M.: *J. Chem. Soc.* **75**, 1003 (1899).  
(119) LETSCHE, E.: *Z. physiol. Chem.* **61**, 215 (1909).  
(120) LIEB, H.: *Mikrochemie* **9**, 105 (1931).  
(121) LIFSCHÜTZ, J.: *Z. physiol. Chem.* **50**, 436 (1906); **91**, 309 (1914).  
(122) MACLEAN, I. S.: *Biochem. J.* **22**, 22 (1928).  
(123) MCCORQUODALE, D. W., THAYER, S. A., AND DOIST, E. A.: *J. Biol. Chem.* **99**, 327 (1933).  
(124) MCCORQUODALE, D. W., LEVIN, L., AND THAYER, S. A.: *J. Biol. Chem.* **105**, lv (1934).  
(125) MAKINO, H.: *Z. physiol. Chem.* **220**, 49 (1933).  
(126) MARRIAN, G. F.: *Biochem. J.* **23**, 1090 (1929).  
(127) MARRIAN, G. F., AND HASLEWOOD, G. A. D.: *Chemistry & Industry* **51**, 277 (1932).  
(128) MARRIAN, G. F.: *Physiol. Rev.* **13**, 185 (1933); *Science Progress* **109**, 69 (1933).  
(129) MARSSON, TH.: *Arch. Pharm.* **58**, 138 (1849).  
(130) MAUTNER, L.: *Monatsh.* **28**, 1113 (1907); **30**, 635 (1909).  
(131) MAZZA, F. P., AND STOLFI, G.: *Chim. ind. agr. biol.* **8**, 408 (1932).  
(132) MENDEL, L. B., AND DANIELS, A. L.: *J. Biol. Chem.* **13**, 71 (1912).  
(133) MENSCHICK, W., PAGE, I. H., AND BOSSERT, K.: *Ann.* **495**, 225 (1932).  
(134) MICHAELIS, L., AND MIZUTANI, M.: *Z. physik. Chem.* **116**, 135 (1925); **118**, 310 (1927).  
(135) MYLIUS, F.: *Ber.* **19**, 369 (1886).  
(135a) MYLIUS, F.: *Ber.* **20**, 1968 (1887).  
(136) NEWMAN, C. E.: *Beitr. path. Anat.* **86**, 187 (1931).  
(137) OKADA, S.: *J. Physiol.* **50**, 114 (1916).  
(138) OTTENBERG, R.: Private communication.  
(139) OTTENBERG, R., AND KAHN, J.: *Proc. Soc. Exptl. Biol. Med.* **29**, 573 (1932).  
(140) PAGE, I. H.: *J. Biol. Chem.* **57**, 471 (1923).  
(141) PAGE, I. H., AND MUELLER, EU.: *Z. physiol. Chem.* **204**, 13 (1931).  
(142) PETTENKOFER, M.: *Ann.* **52**, 90 (1844).  
(143) PICKARD, R. H., AND YATES, J.: *J. Chem. Soc.* **95**, 638 (1908).

- (144) PRINGSHEIM, H.: Ber. **48**, 1324 (1915).  
(145) PRYDE, J.: Nature **133**, 237 (1934).  
(146) QUAGLIARIELLO, G.: Atti accad. Lincei **16**, 387 (1932); International Physiological Congress, Rome, 1932.  
(147) RAUDNITZ, H., PETRŮ, F., AND HAUROWITZ, F.: Z. physiol. Chem. **209**, 103 (1932).  
RAUDNITZ, H., PETRŮ, F., AND STADLER, A.: Ber. **66**, 879 (1933).  
(148) REIMANN, A. K., AND MINOT, A. S.: J. Biol. Chem. **45**, 139 (1921).  
(149) REINDEL, F., AND WEICKMANN, A.: Ann. **478**, 86 (1929).  
(150) REINDEL, F., AND KIPPAN, H.: Ann. **493**, 181 (1932).  
(151) REINHOLD, J. G., AND FERGUSON, L. K.: J. Exptl. Med. **49**, 681 (1929).  
(152) REINHOLD, J. G., AND WILSON, D. W.: J. Biol. Chem. **96**, 637 (1932).  
(153) RHEINBOLDT, H.: J. prakt. Chem. **111**, 242 (1925).  
RHEINBOLDT, H., AND KIRCHEISEN, M.: J. prakt. Chem. **113**, 348 (1926).  
RHEINBOLDT, H.: Ann. **451**, 256 (1926); Z. physiol. Chem. **180**, 180 (1928).  
RHEINBOLDT, H., KÖNIG, O., AND OTTEN, R.: Ann. **473**, 249 (1929).  
(154) RHEINBOLDT, H., KÖNIG, O., AND FLUME, E.: Z. physiol. Chem. **184**, 219 (1929).  
(155) ROSENHEIM, O., AND KING, H.: Chemistry & Industry **51**, 464 (1932); Nature **13**, 315 (1932). The Ring System of Sterols and Bile Acids. Paper No. 1.  
(156) ROSENHEIM, O., AND KING, H.: Chemistry & Industry **51**, 954 (1932). Paper No. 2.  
(157) ROSENHEIM, O., AND KING, H.: Chemistry & Industry **52**, 299 (1933). Paper No. 3.  
(158) ROSENHEIM, O., AND KING, H.: Chemistry & Industry **53**, 91 (1934). Paper No. 4.  
(159) ROSENHEIM, O., AND KING, H.: Chemistry & Industry **53**, 196 (1934). Paper No. 5.  
(160) ROSENHEIM, O., AND STARLING, W. W.: Chemistry & Industry **53**, 1056 (1933).  
(161) ROSENHEIM, O., AND KING, H.: Ann. Rev. Biochem. **3**, 87 (1934).  
(162) ROUS, P., AND MACMASTER, P. D.: J. Exptl. Med. **34**, 47 (1921).  
(163) RUZICKA, L., AND RUDOLPH, E. A.: Helv. Chim. Acta **10**, 916 (1927).  
(164) RUZICKA, L., AND THOMANN, G.: Helv. Chim. Acta **16**, 216 (1933).  
RUZICKA, L., GOLDBERG, M. W., AND THOMANN, G.: Helv. Chim. Acta **16**, 812 (1933).  
(165) SANDQVIST, H., AND GORTON, J.: Ber. **63**, 1935 (1930).  
SANDQVIST, H., AND BENGTSSON, E.: Ber. **64**, 2167 (1931).  
(166) SCHENCK, M.: Zur Kenntnis der Gallensäuren. Paper No. 4, Z. physiol. Chem. **89**, 360 (1914). Paper No. 5, Z. physiol. Chem. **104**, 284 (1919). Paper No. 6, Z. physiol. Chem. **107**, 152 (1919).  
SCHENCK, M., AND KIRCHHOF, H.: Paper No. 16, Z. physiol. Chem. **166**, 142 (1927). Paper No. 32, Z. physiol. Chem. **214**, 42 (1933).  
(166a) SCHENCK, M.: Paper No. 7, Z. physiol. Chem. **110**, 167 (1920).  
(167) SCHENCK, M.: Paper No. 11, Z. physiol. Chem. **139**, 39 (1924).  
(168) SCHMIDT, C. L. A., AND DART, A. E.: J. Biol. Chem. **45**, 415 (1920).  
(169) SCHÖNHEIMER, R., ANDREWS, E., AND HRDINA, L.: Z. physiol. Chem. **208**, 182 (1932).  
(170) SEKITOO, T.: Z. physiol. Chem. **199**, 225 (1931).  
(171) SEYDERHELM, R., TAMMANN, H., AND BAUMANN, W.: Z. exptl. Med. **66**, 539 (1929).

- (172) SHIMIZU, T., AND HATAKEYAMA, T.: *Z. physiol. Chem.* **182**, 57 (1929).  
(173) SHIMIZU, T., ODA, T., AND MAKINO, H.: *Z. physiol. Chem.* **213**, 136 (1932).  
(174) SIGNER, R.: *Ann.* **478**, 246 (1930).  
(175) SOBOTKA, H., AND GOLDBERG, A.: *Biochem. J.* **26**, 555, 557, 566 (1932).  
(176) SOBOTKA, H.: *Naturwissenschaften* **19**, 595 (1931).  
SOBOTKA, H., AND GOLDBERG, A.: *Biochem. J.* **26**, 905 (1932).  
(177) SOBOTKA, H., AND KAHN, J.: *Biochem. J.* **26**, 898 (1932); *Ber.* **65**, 227 (1932).  
(178) SOBOTKA, H.: In Harrow-Sherwin's *Biochemistry*, Chap. IV. Philadelphia (1935).  
(179) STEPHAN, R.: *Münch. med. Wochschr.* **76**, 1579 (1929).  
(180) STOCKHOLM, M., AND SCHMIDT, C. L. A.: *Proc. Soc. Exptl. Biol. Med.* **30**, 877 (1933).  
(181) SZILLARD, P.: *Biochem. Z.* **159**, 325 (1925); **173**, 440 (1926).  
(182) TANRET, C.: *Compt. rend.* **147**, 75 (1908).  
(183) THUDICHUM, J. L. W.: *Ann. Clin. Chem.* **2**, 251 (1881).  
(184) TSCHESCHE, R.: *Z. physiol. Chem.* **203**, 263 (1932).  
(185) TSCHESCHE, R.: *Ann.* **498**, 185 (1932).  
(186) UDRANSKY, L. VON: *Z. physiol. Chem.* **12**, 355 (1888).  
(187) URAKI, Z.: *Z. physiol. Chem.* **207**, 16 (1932).  
(188) VERZAR, F., AND KUTHY, A.: *Biochem. Z.* **205**, 369 (1929); **210**, 265 (1929).  
(189) VERZAR, F.: *Klin. Wochschr.* **12**, 489 (1933).  
(190) VESTERBERG, A.: *Ber.* **36**, 4200 (1903).  
(191) VITALI: *L'Orosi* **3**, 397 (1881).  
(192) VOCKE, F.: *Z. physiol. Chem.* **191**, 83 (1930).  
(193) VORLÄNDER, D.: *Ann.* **345**, 155, 192 (1906).  
(194) WAHLGREN, V.: *Z. physiol. Chem.* **46**, 556 (1902).  
(195) WALLACH, O.: *Ann.* **312**, 171 (1899).  
(196) WALLACH, O.: *Ann.* **331**, 324 (1902).  
(197) WALLACH, O.: *Ann.* **356**, 231 (1907).  
(198) WESTPALEN, TH.: *Ber.* **48**, 1064 (1915). Windaus, Paper No. 22.  
(199) WIELAND, HEINRICH, AND WEIL, F. J.: *Z. physiol. Chem.* **80**, 287 (1912). Untersuchungen über die Gallensäuren. Paper No. 1.  
(200) WIELAND, H., AND SORGE, H.: *Z. physiol. Chem.* **96**, 1 (1916). Paper No. 2.  
(201) WIELAND, H., AND SORGE, H.: *Z. physiol. Chem.* **98**, 59 (1916). Paper No. 3.  
(202) WIELAND, H., AND STENDER, H.: *Z. physiol. Chem.* **106**, 181 (1919). Paper No. 4.  
(203) WIELAND, H., AND BORSCH, E.: *Z. physiol. Chem.* **106**, 190 (1919). Paper No. 5.  
(204) WIELAND, H., AND KULENKAMPPF, A.: *Z. physiol. Chem.* **108**, 295 (1920). Paper No. 6.  
(205) WIELAND, H.: *Z. physiol. Chem.* **108**, 306 (1920). Paper No. 7.  
(206) WIELAND, H., AND WEYLAND, P.: *Z. physiol. Chem.* **110**, 123 (1920). Paper No. 8.  
(207) WIELAND, H., AND SCHULENBURG, W.: *Z. physiol. Chem.* **114**, 167 (1921). Paper No. 10.  
(208) WIELAND, H., AND SCHLICHTING, O.: *Z. physiol. Chem.* **119**, 76 (1922). Paper No. 11.  
(209) WIELAND, H., AND SCHLICHTING, O.: *Z. physiol. Chem.* **120**, 227 (1922). Paper No. 12.  
(210) WIELAND, H., AND ADICKES, F.: *Z. physiol. Chem.* **120**, 232 (1922). Paper No. 13.



- (211) WIELAND, H., AND SCHLICHTING, O.: *Z. physiol. Chem.* **123**, 213 (1922). Paper No. 14.
- (212) WIELAND, H., AND FUKELMAN, L.: *Z. physiol. Chem.* **130**, 144 (1923). Paper No. 16.
- (213) WIELAND, H., HONOLD, E., AND PASCUAL-VILLA, J.: *Z. physiol. Chem.* **130**, 326 (1923). Paper No. 17.
- (214) WIELAND, H., FLASCHENTRÄGER, B., AND REWOLLE, A.: *Z. physiol. Chem.* **134**, 140 (1924). Paper No. 18.
- (215) WIELAND, H., AND MOTHES, W.: *Z. physiol. Chem.* **134**, 149 (1924). Paper No. 19.
- (216) WIELAND, H., AND SCHLICHTING, O.: *Z. physiol. Chem.* **134**, 276 (1924). Paper No. 20.
- (217) WIELAND, H., AND REVEREY, G.: *Z. physiol. Chem.* **140**, 186 (1924). Paper No. 21.
- (218) WIELAND, H., AND JACOBI, R.: *Z. Physiol. Chem.* **143**, 232 (1925). Paper No. 23.
- (219) WIELAND, H., AND SCHLICHTING, O.: *Z. physiol. Chem.* **150**, 267 (1925). Paper No. 24.
- (220) WIELAND, H., SCHLICHTING, O., AND LANGSDORFF, W. VON: *Z. physiol. Chem.* **161**, 74 (1926). Paper No. 25.
- (221) WIELAND, H., SCHLICHTING, O., AND JACOBI, R.: *Z. physiol. Chem.* **161**, 80 (1926). Paper No. 26.
- (222) WIELAND, H.: *Z. physiol. Chem.* **167**, 70 (1927). Paper No. 27.
- (223) WIELAND, H., AND VOCKE, F.: *Z. physiol. Chem.* **177**, 68 (1929). Paper No. 28.
- (224) WIELAND, H., AND VOCKE, F.: *Z. physiol. Chem.* **191**, 69 (1930). Paper No. 30.
- (225) WIELAND, H., ERTHEL, L., AND DANE, E.: *Z. physiol. Chem.* **194**, 107 (1930). Paper No. 31.
- (226) WIELAND, H., AND POSTERNAK, TH.: *Z. physiol. Chem.* **197**, 17 (1931). Paper No. 34.
- (227) WIELAND, H., AND DANE, E.: *Z. physiol. Chem.* **206**, 225 (1932). Paper No. 37.
- (228) WIELAND, H., AND DANE, E.: *Z. physiol. Chem.* **206**, 243 (1932). Paper No. 38.
- (229) WIELAND, H., AND DANE, E.: *Z. physiol. Chem.* **210**, 268 (1932). Paper No. 39.
- (230) WIELAND, H., AND KRAFT, K.: *Z. physiol. Chem.* **211**, 203 (1932). Paper No. 42.
- (231) WIELAND, H., DANE, E., AND SCHOLZ, E.: *Z. physiol. Chem.* **211**, 261 (1932). Paper No. 43.
- (232) WIELAND, H., AND DANE, E.: *Z. physiol. Chem.* **212**, 41 (1932). Paper No. 44.
- (233) WIELAND, H., AND DANE, E.: *Z. physiol. Chem.* **212**, 263 (1932). Paper No. 45.
- (234) WIELAND, H., AND KISEI, S.: *Z. physiol. Chem.* **214**, 47 (1933). Paper No. 47.
- (235) WIELAND, H., AND DANE, E.: *Z. physiol. Chem.* **216**, 91 (1933). Paper No. 50.
- (236) WIELAND, H., AND DANE, E.: *Z. physiol. Chem.* **219**, 240 (1933). Paper No. 52.
- (237) WIELAND, H., AND ALLES, R.: *Ber.* **55**, 1789 (1922).
- (238) WIELAND, H., AND JACOBI, R.: *Ber.* **59**, 2064 (1926).
- (239) WIELAND, H., STRAUB, H., AND DORFMÜLLER, TH.: *Z. physiol. Chem.* **186**, 97 (1930).
- (240) WIELAND, H., AND VOCKE, F.: *Ann.* **481**, 215 (1930).
- (241) WIELAND, H., HESSE, G., AND MEYER, H.: *Ann.* **493**, 272 (1932).
- (242) WIELAND, HERMANN, AND HILDEBRAND, TH.: *Arch. exptl. Path. Pharmacol.* **85**, 199 (1920).
- WIELAND, HERMANN: *Arch. exptl. Path. Pharmacol.* **86**, 79, 92 (1920).

- (243) WILLSTÄTTER, R., AND MEYER, E.: Ber. **41**, 2199 (1908).  
(244) WINDAUS, A.: Ber. **37**, 2027 (1904). Über Cholesterin. Paper No. 2.  
(245) WINDAUS, A.: Ber. **39**, 2008 (1906). Paper No. 6.  
(246) WINDAUS, A.: Ber. **41**, 2558 (1908). Paper No. 11.  
(247) WINDAUS, A.: Ber. **42**, 3770 (1909). Paper No. 12.  
(248) WINDAUS, A., AND RESAU, C.: Ber. **46**, 1246 (1913). Paper No. 16.  
(249) WINDAUS, A., AND UIBRIG, CL.: Ber. **47**, 2384 (1914). Paper No. 19.  
(250) WINDAUS, A., AND UIBRIG, CL.: Ber. **48**, 857 (1915). Paper No. 21.  
(251) WINDAUS, A.: Ber. **49**, 1724 (1916). Paper No. 24.  
(252) WINDAUS, A., AND NEUKIRCHEN, K.: Ber. **52**, 1955 (1919). Paper No. 28.  
(253) WINDAUS, A., AND KIRCHNER, E.: Ber. **53**, 614 (1920). Paper No. 30.  
(254) WINDAUS, A.: Ber. **42**, 238, 616 (1909).  
(255) WINDAUS, A., AND STADEN, A. VON: Z. physiol. Chem. **117**, 147 (1921).  
(256) WINDAUS, A., HÜCKEL, W., AND REVEREY, G.: Ber. **56**, 91 (1923).  
(257) WINDAUS, A., AND BOHNE, A.: Ann. **433**, 278 (1923).  
(258) WINDAUS, A., AND HAUTH, A.: Ber. **39**, 4378 (1906); **40**, 3681 (1907).  
WINDAUS, A., AND BRUNKEN, I.: Z. physiol. Chem. **140**, 47 (1924).  
(259) WINDAUS, A., BOHNE, A., AND SCHWARZKOPF, E.: Z. physiol. Chem. **140**, 177 (1924).  
(260) WINDAUS, A., AND HOSZFELD, R.: Z. physiol. Chem. **145**, 177 (1925).  
(261) WINDAUS, A., AND VAN SCHOOR, A.: Z. physiol. Chem. **148**, 225 (1925).  
(262) WINDAUS, A., AND BOHNE, A.: Ann. **442**, 7 (1925).  
(263) WINDAUS, A.: Ann. **447**, 233 (1926).  
(264) WINDAUS, A., AND VAN SCHOOR, A.: Z. physiol. Chem. **173**, 312 (1928).  
(265) WINDAUS, A., BERGMANN, W., AND KÖNIG, G.: Z. physiol. Chem. **189**, 148 (1930).  
(266) WINDAUS, A., AND TSCHESCHE, R.: Z. physiol. Chem. **190**, 51 (1930).  
(267) WINDAUS, A.: Z. physiol. Chem. **213**, 147 (1932).  
(268) WOLFF: Ann. **394**, 86 (1912).  
(269) LETTRÉ, H.: Ann. **511**, 280 (1934).  
(270) CIAMICIAN, G., AND SILBER, P.: Ber. **41**, 1071 (1908).  
(271) CHEN, K. K., AND CHEN, A. L.: Arch. intern. pharmacodynamie **47**, 297 (1934).  
(272) FELS, E., AND SLOTTA, K. H.: Klin. Wochschr. **10**, 1639 (1931).  
ALLEN, W. M.: J. Biol. Chem. **98**, 591 (1932).  
FEVOLD, H. L., AND HISAW, F. L.: Proc. Soc. Exptl. Biol. Med. **29**, 620 (1932).  
(273) BUTENANDT, A., WESTPHAL, U., ET AL.: Z. physiol. Chem. **227**, 84 (1934); Ber. **67**, 1440, 1893, 1897, 2088, 2092 (1934).  
WINTERSTEINER, O., AND ALLEN, W. M.: J. Biol. Chem. **107**, 321 (1934).  
SLOTTA, K. H., RUSCHIG, H., AND FELS, E.: Ber. **67**, 1270 (1934).  
HARTMANN, M., AND WETTSTEIN, A.: Helv. Chim. Acta **17**, 878 (1934).  
(274) BUTENANDT, A., AND SCHMIDT, J.: Ber. **67**, 1901 (1934).  
(275) BUTENANDT, A., WESTPHAL, U., AND COBLER, H.: Ber. **67**, 1611, 2085 (1934).  
FERNHOLZ, E., AND CHAKRAVORTY, P. N.: Ber. **67**, 1855, 2021, 2027 (1934).  
(276) BUTENANDT, A. AND TSCHERNING, H.: Z. physiol. Chem. **229**, 167, 185, 192 (1934).  
(277) RUZICKA, L., GOLDBERG, M. W., ET AL.: Helv. Chim. Acta **17**, 1389, 1395 (1934).  
(278) CHARGAFF, E., AND ABEL, G.: Biochem. J. **28**, 1901 (1934).  
(279) WINTERSTEIN, A., AND SCHÖN, K.: Naturwissenschaften **22**, 237-8 (1934).



# THE PHYSICAL CHEMISTRY OF FOAMS<sup>1</sup>

SOPHIA BERKMAN AND GUSTAV EGLOFF

*Research Laboratories, Universal Oil Products Co., Chicago, Illinois*

*Received August 25, 1934*

## CONTENTS

I. Introduction.....	378
II. Foam formation.....	378
A. The origin and mechanism of foam formation.....	378
1. The formation of a single lamina.....	378
2. Forces acting in a lamina.....	381
3. Construction of a foam system.....	381
4. Conclusions.....	382
B. Factors influencing foam formation.....	383
1. Vapor pressure and surface tension.....	383
2. Solubility and surface activity.....	386
3. Solubility, spreading, and viscosity.....	387
4. Boundary heterogeneity.....	388
5. Conclusions.....	394
III. The stability of foam systems.....	394
A. Factors in the stabilization of foam systems.....	397
1. The influence of electrolytes.....	398
2. The action of protective colloids.....	400
3. Viscosity as a factor in the stabilization of foam systems.....	400
4. The action of finely dispersed solids.....	402
B. Forces involved in maintaining stability.....	403
1. Electrostatic and capillary forces.....	403
2. The adsorption layer in the interface of a foam system and its thickness as a factor influencing stability; the Langmuir-Harkins orientation theory.....	405
C. Conclusions.....	409
IV. Transition to the unstable state.....	410
A. Forces acting in the thinning of a film.....	410
B. Forces influencing transition to the unstable state.....	412
C. Conclusions.....	416
V. Foam prevention.....	417
A. Factors influencing foam prevention.....	417
B. Examples of desirable and undesirable foaming.....	419
VI. Summary.....	420

---

<sup>1</sup> This paper was presented before the Division of Physical and Inorganic Chemistry at the Eighty-fifth Meeting of the American Chemical Society held in Washington, D. C., March 24, 1933.

## I. INTRODUCTION

A parallel exists between the conception of emulsions and the conception of foams. Not only are foams similar to emulsions, but they may be considered, to a certain extent, as concentrated emulsions having a gas instead of a liquid as the dispersed component. On the other hand, the fact that the formation of an emulsion is often accompanied by the formation of foam indicates that at least in these cases the conditions under which foam systems originate are identical with those of emulsions. Oils, readily emulsified in water, act as foam-forming substances, while difficultly emulsifiable oils, such as cedar or castor oil, show very little tendency to form foam. Foams, as well as emulsions, are generally three-component systems, that is, they require the presence of an agent for their formation. Factors influencing the origin of foam systems and those influencing their ability to sustain themselves indicate another similarity between the properties of emulsion and foam systems.

Three stages may be differentiated in the life of dispersion systems such as emulsions and foams: (1) origin; (2) maintenance of the stability of the system formed; and (3) transition to the unstable state.

Factors and forces governing different stages in the existence of a foam comprise the physical chemistry of the system in question.

## II. FOAM FORMATION

### A. THE ORIGIN AND MECHANISM OF FOAM FORMATION

#### *1. The formation of a single lamina*

Foam is a formation in which macroscopic, microscopic and/or ultra-microscopic gas bubbles are separated by micronic or submicronic liquid layers (skins). Therefore, foam is a union of laminae. The formation of a foam occurs in the same manner as the formation of a single liquid lamina. Each film formed serves as a basis for the formation of other surface films, so that knowledge concerned with the formation of a single film may account to some extent for a number of films called in the aggregate "foam." The construction of a foam system may be described as follows. A bubble of air rising in a liquid reaches the surface but does not break through; it rather lifts up a film in the form of a spherical wedge. A second bubble rises to the top, approaches the surface at a certain point, and is placed under the first film. This determines the formation of a second lamina which lifts the first film on one side so that the amount of air enclosed by these two laminae respectively is separated by a portion of the second lamina, as well as by a partition of liquid. In the event the second lamina originates at a distance, the spherical wedge of the first is then completed, the capillary forces and flotation of light bodies assisting the second bubble

to approach the first bubble in order to create a tangential contact provided they are not separated by too great a distance, which is not the case in foam formation.

Actually, only a part of a bubble reaches the surface and wets it. The faces of the bubble change in curvature before they touch tangentially, and the same change occurs also along the line on which a bubble is bound to other bubbles. Obviously, because of their liquid nature, the two laminae cannot meet each other either at their angles or at their linear edges; for continuity, it is necessary that they be formed along the line of meeting in a small area of concave surfaces placed perpendicularly to this line. But this area is too small to be distinguished, as are the partitions, which unite to form a small mass having its own curvature independent of the respective curvatures of the two laminae and dependent on the difference in the pressure exerted upon the sides of the two volumes of air enclosed. If these two volumes of air are equal, the two laminae belong to equal spheres, compressing the two volumes of air with the same intensity. But if the volumes of air are unequal, the two laminae belong to spheres of different diameters, exerting an unequal pressure on the two volumes of air. Consequently the partition, when subjected to equal pressure on both sides, is not curved, but plane, and when subjected to unequal pressure, is distorted on one side, the partition diminishing until the pressure produced by it in the concave curvature counterbalances the elasticity of the air.

Since air bubbles are always rising, the structure formed is lifted continuously. All laminae are layered with more or less symmetry, depending upon the difference in the volume of air bubbles and the distribution of points where they may reach the surface of the liquid, the air bubble structure growing higher and higher.

Observation reveals that this accumulation of bubbles is not governed by chance, but is subject to law, that is, innumerable bubbles join together in threes at equal angles in such a way that not more than four edges of three laminae pass through the same point, cutting laminae, as well as edges, at an angle of  $120^\circ$ . Therefore, foam bubbles, in cross section, are hexagonal structures with equal angles of  $120^\circ$  (44).

In figure 1 a configuration of a group of four bubbles, assumed to be equal in diameter and combined rectangularly, is graphically presented. But since this structure is in a state of temporary equilibrium, a transition from this structure to a structure in which three laminae are joined together at an angle of  $120^\circ$  may occur. In this same figure there is also included a projection of assembled curved laminae connected with one another by other laminae extending the length of the prism.

In the formation of a foam, bubbles are formed in which more than three laminae come together, the laminae sliding over one another until condi-

tions for true equilibrium are fulfilled. The motion of the liquid, when displaced by a bubble, lifts up the skin formed and is such that it stretches the two surfaces where the liquid comes in contact with the air. The liquid film or lamina may be stretched only when surface tension and vapor pres-

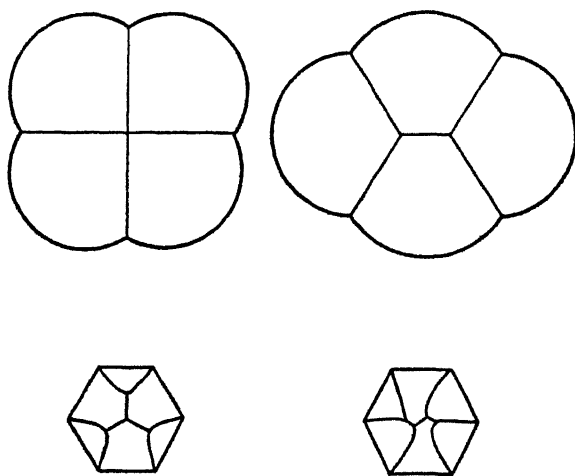


FIG. 1. A GROUP OF FOUR BUBBLES AND A PROJECTION OF ASSEMBLED CURVED LAMINAE

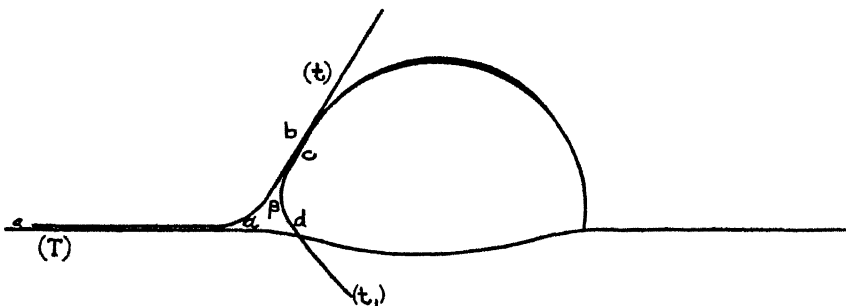


FIG. 2. GRAPHIC PRESENTATION OF A LAMINA RESTING ON THE LIQUID FROM WHICH IT ORIGINATED

sure are small. In 1869, Mensbrugghe (37) proposed a formula for the equilibrium of a lamina resting on the same liquid from which it originated (refer to figure 2 for graphic presentation).

In figure 2 the interior surface of a bubble is concave.

## 2. Forces acting in a lamina

The surface tension,  $T$ , of the external liquid at  $a$ , the horizontal components of two surface tensions  $t$  at  $b$  and  $c$ , and the horizontal component of a single surface tension  $t'$  at  $d$  are the forces which tend to displace the liquid contained in the area bounded by curves  $ab$  and  $cd$  and line  $ad$  (figure 2). (These tensions are all directed against the exterior of the small mass of the liquid.) If  $\alpha$  and  $\beta$  represent angles formed by tangents with the horizontal in  $b$  or  $c$  and in  $d$ , then we have:

$$T = 2t \cos \alpha + t' \cos \beta$$

According to Mensbrugghe if, in the case of an aqueous soap lamina formed on the surface of the same liquid from which it originated, tensions  $T$ ,  $t$ , and  $t'$  are equal to each other, the equation is then written:

$$2\cos \alpha + \cos \beta = 1$$

Or, if the lamina is very large, the pressure produced upon the air enclosed is very low; therefore,  $\beta$  is close to zero,  $\alpha$  close to  $90^\circ$ , and the lamina hemispherical.

If the dimensions of the sphere decrease, the pressure of the air on the interior, as well as  $\beta$ , increases with a gradual decrease in  $\alpha$ ; the lamina would then constitute a portion progressively smaller than a complete sphere. Also, it is evident that if  $\cos \alpha$  is close to 1, then  $\cos \beta$  is close to  $-1$ . This is in the case of an extremely small air bubble; in the surface of the liquid, practically no bubbles originate.

## 3. Construction of a foam system

In passing from a single liquid lamina to a system made up of a number of laminae, it is of great importance to consider the surface tension factor. Quincke (46) and Marangoni (35) studied the conditions under which a liquid spreads into a thin film, concluding that: (1) Surface tension exists in all liquid surfaces and therefore in all films and is independent of the curvature of the surface or film. This surface tension is the same for all directions of the surface or film and at every point. (2) Surface tension varies with the nature of the liquid. (3) For the same liquid, surface tension varies inversely with the temperature, but at ordinary temperature undergoes little change.

Many hypotheses have been proposed to account for surface tension: according to Segner (56) surface tension originates from the mutual attraction of molecules in the superficial layer tangential to the curvature of this layer; according to Mossotti (41), Hough (26), and Dupré (12), surface tension depends upon the asymmetry of molecular action in the



thickness of the superficial layer, although only minor differences in this type of asymmetry are involved; Hagen's (17) hypothesis attributes surface tension to a great density of the superficial layer; the hypotheses of Lamarle (29) and Henry (23) consider surface tension to be due to the contraction of the superficial layers, one portion of the molecules in this layer passing into the interior, thereby decreasing spreading.

According to the work of Laplace, who considered only the effect of curvature, the pressure from a concave layer in a state of equilibrium opposes a force originating in general hydrostatic actions. It must be borne in mind that convex and concave superficial layers exert a normal pressure on a liquid at each point, the superficial layer, meeting an opposing force, finally spreading after curving beyond a critical point. In other words, molecules in the same layer must be in a condition to spread tangentially; this is the reverse of Young's theory, which assumes surface tension and shows that the tangential component of the surface tension around a given point has a normal pressure as a resultant independent of whether the surface is convex or concave. This pressure, as Laplace showed, exists only because of molecular attraction and may be analyzed around a point into tangential components, each of which constitutes a surface tension.

Plateau states that surface tension is a result of curvature, but is independent of the value of this curvature. From an actual consideration of the two tangential components of a convex surface of a liquid and the initial pressure which causes their appearance, it is evident that if the curvature decreases this pressure decreases at the same time but, on the other hand, the angle between the two components increases in such a manner that these components or surface tensions retain the same value as for the previous curvature. If the curvature decreases up to a point where it disappears, the component effected by pressure is eliminated and the two remaining components become prolongations of each other; therefore curvature is infinite, as in the case of plane liquid surfaces.

Another question to be considered is how surface tension penetrates similar surfaces, for instance, a film of liquid adhering by its edges to the surface of a liquid, or to other films. In the case of films, these partitions are invisible, but their existence is obvious. As previously stated, the faces of the bubbles change their curvature before they touch tangentially.

#### *4. Conclusions*

1. Foam is a union of laminae.
2. The mechanics of foam formation is similar to that of a single lamina.
3. The symmetry in the construction of a foam depends upon the difference in the volume of gas enclosed by a lamina.
4. The accumulation of bubbles in a foam follows Plateau's rule, that is, bubbles join in three at angles of  $120^\circ$ .

5. Physical factors involved in the spreading of a liquid into a thin film apply to foams, surface tension being one of them. Surface tension is the result of curvature and independent of the value of curvature. Surface tension is the same at every point in all tangential directions. Surface tension varies with the nature of the liquid and inversely with temperature. The change in curvature of the faces of bubbles prior to their tangential contact is proof that surface tension penetrates a film.

## B. FACTORS INFLUENCING FOAM FORMATION

### 1. *Vapor pressure and surface tension*

Not every gas bubble distribution in a liquid, as in a sol, can be designated a foam. Foam formation combined with a great increase in surface is possible only when the surface tension of the foam components is small.

Pure liquids with low surface tension are close to their boiling points and have a high vapor pressure. Therefore, thin liquid laminae of which foam is composed evaporate too quickly to form and sustain foam. On the other hand, when liquids have a low vapor pressure, their surface tension is too great. Low surface tension and vapor pressure when present together in capillary-active substances serve as favorable conditions for the promotion of the foaming process (15). Capillary-active substances, when dissolved in a system, lower the surface tension and are adsorbed in the interface.

That a certain relation between foaming and surface tension of the liquid from which foam originates should be anticipated follows from the fact that solutions containing dissolved capillary-inactive substances not influencing surface tension do not produce foam. Foams do readily occur, however, with solutions of surface-active substances, such as organic salts and acids, sugars, glycerins, alcohols, and esters. Of course, surface-active substances are foam agents only under definite conditions. In certain cases, discussed later, other factors such as concentration, solubility, viscosity, et al., may overcome the influence of capillary activity. Bartsch (3) proved experimentally (figure 3) that foam formation and lowering of surface tension are antagonistic when a certain concentration is exceeded.

Foam formation-concentration curves drawn by Bartsch for butyric, nonylic, and caproic acids and for octyl, heptyl, and amyl alcohols indicate that these capillary-active foam-forming substances lose their foaming ability when a certain saturation is exceeded. The upper curve b/c of butyric acid has a shape indicating that the relationship between foaming ability and surface tension is lost in higher concentrations. The same experimenter also found that even in the case of such strong capillary-active substances as sodium cholate or saponin the foam-producing ability may be

lost when supersaturation occurs. Wo. Ostwald and A. Steiner (43) proved with humus sols that a relation does not always exist between surface tension and foam-forming ability. Bartsch pointed out that oleic acid emulsions which do not foam show a pronounced foaming ability in certain concentration intervals, while the latter, as a rule, have a surface tension of an order of magnitude corresponding to water. By comparing

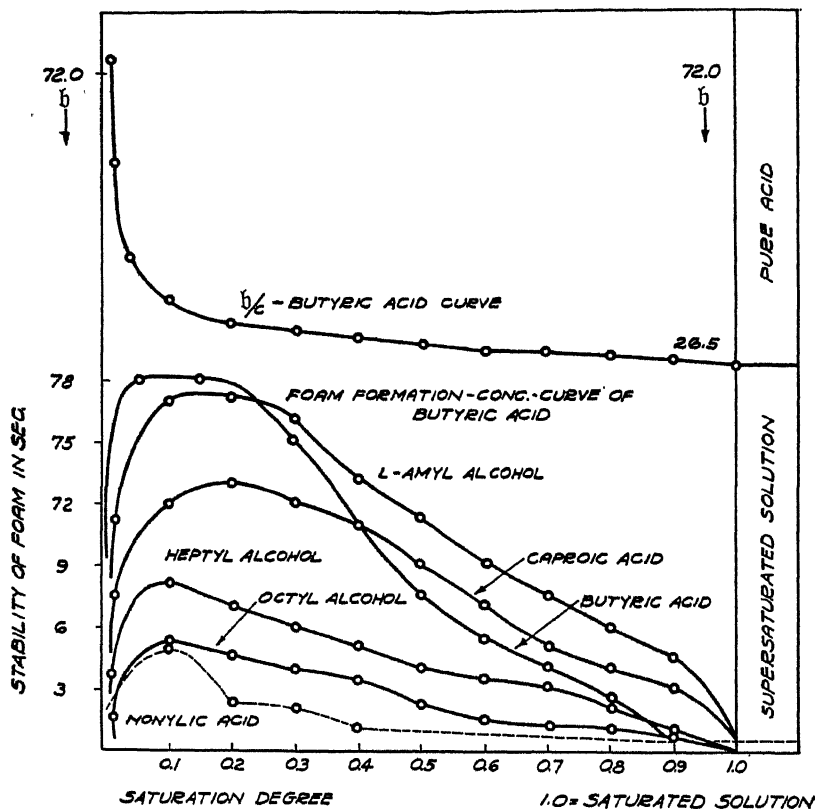


FIG. 3. SURFACE TENSION AND FOAM FORMATION

the foam-forming ability of various surface-active foam-forming substances, such as homologous alcohols and fatty acids, the relationship between surface activity and foam formation was found to be complex. A parallel was observed when concentrations were studied corresponding to the limited value for foam formation (3) (figures 4 and 5, from reference 3).

When the maximum in foam formation was reached, the acids decreased the foaming ability in spite of an increase in surface activity, using members

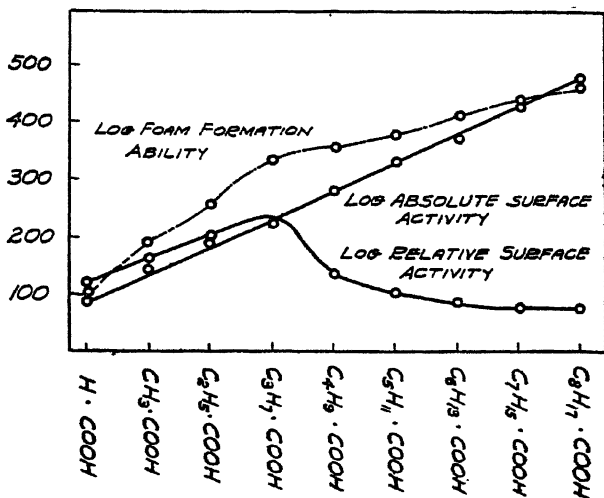


FIG. 4. SURFACE ACTIVITY AND FOAM FORMATION FOR MORE DILUTE SOLUTIONS

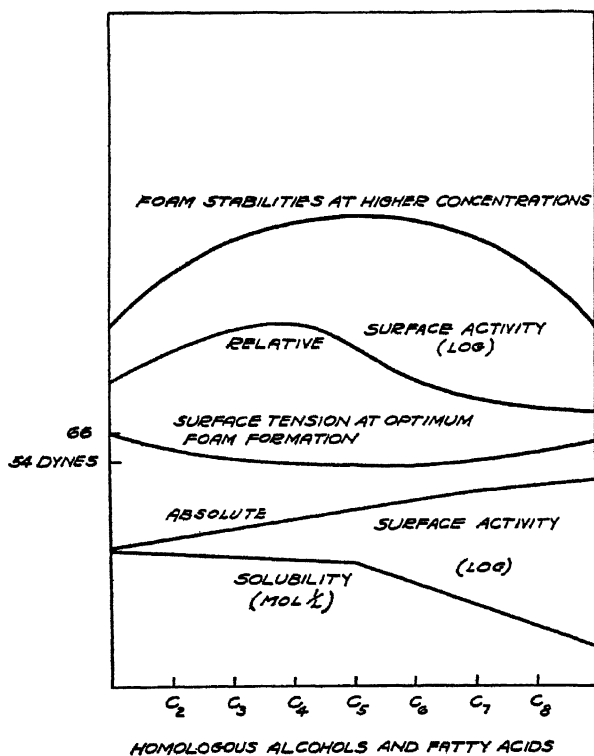


FIG. 5. SCHEME OF THE RELATIONSHIP BETWEEN FOAM FORMATION AND SURFACE ACTIVITY FOR SOLUTIONS OF HIGHER CONCENTRATION

of the homologous series with limited solubility. The only regularity observed by Bartsch was in the group arrangement of surface tension values. The average corresponded to an optimum foam formation at 61 dynes for fatty acids and alcohols. In other cases, the decrease in foaming ability was parallel with the decrease in surface activity of the component dissolved in water. A certain relationship between surface activity and foam formation suggested itself when the solubility of single members of the homologous series was considered. Figure 5, taken from Bartsch's work, shows that foaming ability in very dilute fatty acid and alcohol solutions increases steadily with increasing surface activity.

## 2. Solubility and surface activity

When surface-active foam-forming substances are dissolved partially or completely in molecular dispersion, solubility of the molecularly dispersed substance also plays a certain part in producing foam. Traube and Klein (67) state that only the molecularly dispersed portion of the solute is related to surface activity, therefore the solubility factor must also be considered in foaming. Solubility decreases with increase in the number of carbon atoms; thus the higher members of the series have a limited solubility or are insoluble. On the other hand, Traube's rule indicates that capillary activity increases in ascending the homologous series. The decrease in solubility limits, therefore, the increase in surface activity. This fact led Bartsch to assume that the maximum solubility is reached by the lowest members of the homologous series and the maximum capillary activity by the highest members. Therefore, a favorable relationship should exist between solubility, surface activity, and foaming ability for the respective intermediate members, such as butyric acid or amyl alcohol. In the case of higher members, this relationship appears to be more unfavorable because the solubility decreases more readily than the surface activity increases.

This assumption may be proved by taking into consideration the well-known equation of Szyszkowski, who formulated an empirical expression of the relationship between  $\sigma$  (surface tension) and concentration ( $c$ )

$$\Delta = \frac{\sigma_M - \sigma_L}{\sigma_M} = b \ln \left( \frac{c}{k} + 1 \right)$$

where  $\sigma_L$  denotes the surface tension of the solution,  $\sigma_M$  denotes the surface tension of the solvent, and  $k$  is a constant.

By means of Szyszkowski's (61) formula, values for surface activity relating to solubility are obtained when the constant  $k$ , the reciprocal value of which is a measure for capillary activity, is divided by the amount of surface-active substance in a given volume of solution. The greater the surface activity and the solubility, the smaller the value of the quotient.

The influence of solubility is also quite considerable; thus an increase in the solubility of a substance forming the foam leads to an increase in foaming (3).

In case a second agent is added to a solution of a capillary-active substance, the foaming ability of the mixture may be improved depending on the changes occurring in the solubility. If the solubility of the first surface-active substance is increased by the addition of the second, then the foaming ability of the mixture may be greater than that of either of the components. For example, nonylic acid, a weak foam agent, becomes an excellent foam agent in a 60 per cent aqueous ethyl alcohol solution. Other foam agents may bring about a decrease in foam formation. In the following examples it was shown that a decrease in solubility brought about by a second capillary-active substance produced a decrease in foaming. A solution of 0.5 per cent of isoamyl alcohol and 0.1 per cent of benzyl alcohol decreased the duration of the foam from 17 seconds to 9 seconds; a *m*-cresol solution and 0.01 per cent of benzyl alcohol decreased the duration of the foam from 11 seconds to 6 seconds. A maximum foaming was obtained with 29 per cent ethyl alcohol; larger amounts caused a diminution in foaming.

### *3. Solubility, spreading, and viscosity*

There is another proof for the relationship existing between capillary activity, solubility, and foam formation. Since the solubility of a substance depends on the difference between the attraction of the solute for the solvent and the solute for the solute, it is probable that the work of adhesion minus the work of cohesion, called "spreading coefficient," should also be considered an important factor in determining solubility. But there is a distinction between "spreading in water" and "solubility." In the case of spreading, it is the most active or polar part which is chiefly involved, while the whole molecule relates to solubility. From this standpoint, it is to be inferred that liquids incapable of spreading upon water are insoluble in water. Therefore, the spreading-solubility relationship should be considered in choosing foam agents.

The classical theories (44) ascribe the property of film formation and therefore the property of foam formation to two general factors: (1) viscosity, and (2) low surface tension of the liquid from which foam originates. Since many liquids possessing high viscosity and low surface tension do not form sustained films or foams, many mobile aqueous solutions with a surface tension nearly equal to that of pure water are capable of foaming (36). This assumption has proved unsatisfactory. Tickell (65) observed that gas bubbles liberated from oil upon release of pressure formed a foam. An impure substance containing capillary-active and

emulsifying constituents having low surface tension and high viscosity fulfills the requirements for foam formation, according to this investigator. Tickell assumed that low surface tension is necessary for the development of a broad foam surface. Since the capillary-active constituents of an impure complex substance, such as petroleum, may concentrate in the surface films, this condition is satisfied with oils.

The question concerning the distribution of a substance between a solvent and its surface layer has been discussed by Thomson (64) and Ramsden (48). Zawidski (73) measured aqueous acid solutions to which a sufficient amount of saponin had been added to insure a stable foam on shaking. He proved that the composition of the solution differed from its surface layer, the concentration of the saponin in the surface layer being greater than that in the interior of the solution. This fact agrees with the viscosity and tension relation of aqueous saponin bubbles investigated by Plateau. Under similar experimental conditions, the ratio of concentration of saponin in the foam to that in the solution is not constant. When air is led through an aqueous saponin solution and the foam formed is removed after a certain interval of time, the solution ceases to produce foam, while the foaming ability is definitely increased in a condensed foam; therefore, it is apparent that all the saponin is carried away by the foam. Zawidski ascertained this phenomenon by refractometric measurements. He also observed a second phenomenon, namely, that when alcohol is added gradually to an aqueous solution of saponin a certain point is reached at which the ability to form foam on shaking is eliminated. Zawidski explained this phenomenon by stating that alcohol diminishes the surface tension of water much more than that of saponin. This phenomenon of eliminating the ability to form foam was also observed when solutions of isobutyric acid were added to water, thus greatly diminishing its surface tension. Even a large addition of saponin did not produce a stable foam in this case, but a stable foam was obtained by neutralizing the acid with a base.

On the other hand, it was found that if a capillary-active substance was added to a certain foaming liquid like soap, its ability to foam was maintained if the capillary-active substance had the property of forming surface films or skins.

#### *4. Boundary heterogeneity*

Shorter (59) ascribed the foaming ability of a liquid to the existence of a surface pellicle which was considered as being formed from the dissolved substance. This most fruitful idea of the heterogeneity of the surface layers of a solution was treated theoretically by Gibbs. He proved from thermodynamical considerations that if the addition of a solute lowered

the surface tension of the solvent, the solute existed in a higher degree of concentration in the surface layers than in the solution. The decrease in the surface tension is nothing more than positive adsorption and the excess concentration in the surface towards the solution:

$$x = - \frac{c}{RT} \frac{d\sigma}{dc}$$

where  $c$  denotes the concentration of the solution,  $R$  and  $T$  are constants, and  $d\sigma$  is the surface tension decrease.

This deduction of Gibbs was verified by a number of investigators. Lewis (34) measured the surface excess of various solutes in the case of the interface between various solutions and hydrocarbon oils and found in nearly all cases that the measured excess was very much larger than the theoretical value. Thus in the case of a solution of sodium glycocholate the measured surface excess was found to be  $5.4 \times 10^{-6}$  g. per square centimeter, while the theoretical value was  $7.0 \times 10^{-8}$  g. per square centimeter. Shorter explained this discrepancy as follows: Gibbs' theory is applicable to cases in which the process of formation of the surface layer is a simple reversible effect, while Shorter's experiments on surface layers formed by solutions of saponin, peptones, and albumin showed that the process is thermodynamically irreversible and therefore outside of the scope of Gibbs' theory. Bartsch's investigations showed that the maximum of foam-forming surface-active substances depends considerably on the abrupt drop in the concentration, and that this does not apply to skin-forming substances. Surface tension-concentration curves of aqueous solutions for surface-active substances revealed that the foam-forming concentration coincides with the interval of the steep part of the curve and, conversely, the steep part of the curve indicates that a concentration excess in the Gibbs layer is relatively large for these concentration intervals. This maximum for foaming ability is reached for aqueous surface-active substances when the concentration difference in the boundary layer possesses a maximum value. In case of skin-forming substances the concentration excess in the Gibbs layer, as a result of gel formation, is considerable and the concentration drop between the layer and the solution is very steep.

Regarding the structure of the Gibbs layer at the maximum of foam formation, two cases deserve consideration: (1) When concentration is infinitely small, the Gibbs layer is composed of water molecules. (2) When concentration is infinitely large, i.e., saturated, the Gibbs layer consists of molecules of surface-active substances.

These two extremes, in which the boundary layer is homogeneous with reference to the types of molecules, do not induce foaming. In all intermediate cases, in which this layer is made up of two types of molecules and



is therefore heterogeneous, foaming occurs. The foaming ability of very dilute or very concentrated solutions having homogeneous boundary layers is small because in the first, water molecules, and in the second, surface-active molecules, predominate. Thus the foaming ability increases when the boundary layer becomes heterogeneous (3). From this standpoint, the addition of a second surface-active substance makes the foam less stable. When this addition causes an increase in surface-active molecules in the boundary layer due to the presence of water molecules, the second substance increases the homogeneity in relation to the molecules of the surface-active substance already present and as a result of which the foaming ability decreases.

The decrease in foaming in the case of high molecular members of the homologous series with an increasing tendency for colloidal solubility is explained by its action contrary to heterogeneity due to the formation of large homogeneous molecular aggregates. The molecularly dispersed members of the series have a greater heterogeneity of the Gibbs layer than the corresponding members partially colloiddally dissolved. From this viewpoint, solutions of surface-active substances foam better the more abrupt the concentration drop between the Gibbs layer and the solution, as mentioned previously, and the more heterogeneous the Gibbs layer involved (59). High surface activity, low molecular solubility, and homogeneous boundary layers favor abruptness in the concentration drop (3). Only in intermediate members of the homologous series do both conditions exist.

With respect to colloiddally soluble organic liquids, like oleic acid, turpentine, etc., it is assumed that only those capable of spreading on a water surface act as foam formers. The velocity of spreading is not parallel to the rate in surface increase in the formation of homogeneous layers, therefore there is no possibility for the formation of homogeneous layers, and the foam walls consist of heterogeneous boundary layers (molecules of water and organic substances).

Iliin (27), discussing adsorption and superficial energy at the boundary of various phases, points out that a substance dissolved in water is called active when a slight change in its concentration ( $c$ ) greatly changes the surface tension at the air-solution boundary and that the measure of the activity of the dissolved substance taken in the sense of surface tension (surface activity) is expressed by the equation:

$$g = \frac{\partial \sigma}{\partial c}$$

Neither the activity of an adsorbent, nor the activity of a dissolved substance is involved in Iliin's conception, but only the mutual phase

activity of two or more phases at their boundary. As activity, the gradient of certain energy effects, depending upon the change of one of the variables, especially concentration, is envisioned. Consequently, a gas which has a relatively small adsorption heat and adsorption activity at the same time shows great change in  $\sigma$ ,—a great surface activity as a result of the difference between surface tension of the gas and the adsorbent. Ramsden (50) indicates that adsorbed molecules in a foam are held in direct contact with both gas and water. There is evidence that before adsorption becomes too great every molecule of the foam agent entering into the surface region is forced to the actual interface and, although still soluble, is retained there permanently.

Shorter (58), in his article "On a Classification of Foaming Solutions," states that in order for solutions to be able to form foams they should have a surface tension capable of local variations. This power of local variations in surface tension is due to the existence of a pellicle called the Gibbs layer. Differences in the foaming nature of various solutions considered important by Shorter are not the same as those pointed out earlier by Plateau. Shorter's classification of foam solutions is based upon two distinct points to be considered with respect to the foaming phenomenon: (1) the nature of the process of the formation of the surface layer, and (2) the manner in which the surface layer contributes to the stability of the liquid film.

The process of the formation of the surface layer varies in different solutions. In the case of a mixture of acetic acid and water, equilibrium between surface layers and the solution itself is obtained readily, except in very dilute solutions, so that the process of the formation of the surface layer may be thermodynamically reversible (10). In soap and saponin solutions, the process assumes another nature. Millner (40) observed in solutions of sodium oleate an irreversibility of the formation of the surface layer. In dealing with solutions of saponin, peptone, and albumin, Shorter (57) showed that the formation of a surface layer proceeds for several weeks without indication of an equilibrium. The substance forming the surface layer separates from the body of the solution at a rate independent of the thickness of the surface layer and continues to separate for some time without any sign of equilibrium. This process is thermodynamically irreversible, but in albumin solutions, it appears to be "chemically" irreversible, the film substance being insoluble in water (49).

Shorter classifies foaming solutions into three distinct classes having the following characteristics: (1) Surface concentration thermodynamically reversible; superficial viscosity or rigidity absent. (2) Surface concentration thermodynamically irreversible; superficial viscosity or rigidity absent. (3) Surface concentration thermodynamically irreversible; surface layers especially viscous or rigid.

In case the foam-producing agent is a solid body, the spreading and wetting factors are of significance. Substances less wetted by water adsorb capillary-active substances from aqueous solutions more readily when building the Gibbs layer. Hydrophobic galena, copper pyrites, and zinc blende are not readily wetted, and therefore pass into foam. Quartz and feldspar are hydrophilic and, consequently, are well wetted; therefore they remain in the water when separated from their ores in the flotation process.

Substances like sulfides, insoluble oxides, carbonates, carbon, insoluble sulfates, and metals, in fact all substances, as shown by Reinders (53), which attach themselves to the boundary by distribution between two liquids, or are completely surrounded by a liquid, may be used as foam agents. Reinders and Hoffman (54) investigated adsorption through adhesion at the liquid-gas interface; the theory of these processes was developed by De Coudres (9). When two liquids come in contact with a solid body, then one of these liquids may displace the other and surround it completely; alternately, each liquid may surround the solid partially with the formation of an edge angle at the contact surface.

In the first case we have a complete wetting by one of the two liquids, while in the second case only a partial wetting takes place, the latter occurring when an aqueous suspension of a solid substance is mixed with a second liquid insoluble in water as in the case of water-oil emulsions. When liquid-gas systems, instead of liquids insoluble in water, are dealt with, the gas phase is present. When water displaces the gas phase from the surface of the solid particles, the latter are completely wetted by water and no precipitation occurs. When only a partial displacement of the gas phase occurs with the formation of an edge angle, then the solid particles are attracted to the liquid-gas interface and the aqueous surface layer floats freely. Valentiner (69) actually showed that a powder may float better the larger its edge angle.

Relationships become more complicated when a molecularly or a colloiddally dispersed solution comes in contact with gas in the presence of a surface of solid particles, such as a foam-producing agent or a stabilizer.

Surface tensions at phase boundaries lose their characteristic fixed values, which cause homogeneous wetting ability, and become functions of the concentration of the dissolved substance. The relation between adsorption and wetting ability was also investigated by Traube and Kicke (66), who found that substances which are not readily wetted by water adsorb caprylic acid from aqueous solutions more readily than substances which are more readily wetted by water. The solid phase may also lose its foam-forming ability when a substance insoluble in water and which wets the solid phase more readily than does water itself (chloroform, for example) is added. Further investigations of Traube and Nishizawa (68) on the

adsorption of caprylic acid on minerals showed a relation existing between wetting ability, adsorption, and flotation.

The process of separating particles of finely ground ore from the particles of gangue by flotation is an application of surface phenomena. This process is based on the fact that mineral sulfides, e.g., galena, are wetted by certain oily liquids, but not by water, whereas particles of quartz are wetted by water and not by other liquids. Therefore, if the finely divided ore is agitated with an oil-water froth, the sulfide particles are preferably wetted by the oil films surrounding the bubbles of the froth, while the quartz remains in the water phase. The sulfide particles which attach themselves to the bubbles are floated to the surface by a stream of air and separated from the gangue particles which remain at the bottom of the container. Harkins and Feldman (14) state that since the free surface energy of almost all inorganic solids is high, their work of cohesion is of the same order, and the work of adhesion is also high with reference to practically all liquid substances. Since the work of cohesion in water and organic liquids is generally low, the values of the coefficient of spreading of these liquids upon such solids should generally be high. Therefore the spreading of these liquids should occur upon such solids when their surfaces are pure.

Among other factors, the spreading ability of liquids forming foams indirectly plays an essential part in determining foam formation. Briefly, these views of the criterion of spreading of one liquid upon the other are as follows:

1. According to Lord Rayleigh (52), all liquids spread on pure water.
2. Liquid *b* spreads on liquid *a* if  $T_a$  is greater than  $T_b - T_{ab}$ , where  $T_{ab}$  represents the interfacial tension between two liquids and  $T_a$  and  $T_b$  the respective surface tensions. Non-spreading conditions exist when  $T_a$  is less than  $T_b - T_{ab}$ .
3. According to Langmuir, liquids whose molecules are polar or contain polar groups spread on water.
4. According to Harkins, a liquid spreads if its work of surface cohesion  $W_c$  is less, and will not spread if its work of surface cohesion  $W_c$  is greater than its work of adhesion  $W_a$  with respect to the liquid upon which the spreading is to occur.

Harkins (19) developed a spreading coefficient by thermodynamic reasoning, using two hypotheses as a basis. Only large scale motion is of importance in spreading and only free surface energies are involved. The free energy decrease is  $S = \gamma_a - (\gamma_b + \gamma_{ab})$ , where  $\gamma_{ab}$  represents the free energy of the interface. The work of adhesion  $W_a$  is given by the equation of Dupré as  $W_a = \gamma_a + \gamma_b = \gamma_{ab}$ . The work of cohesion is  $W_c = 2\gamma_b$ . A combination of three equations  $S = W_a - W_c$ , indicates that spreading

occurs if the adhesion between two liquids is greater than the cohesion in the spreading liquid. Thus positive values of  $S$  correspond to spreading and negative values to non-spreading. In general, Harkins assumes that the value of the coefficient increases as the electromagnetic field around the molecules becomes more asymmetrical. This is due to the fact that in the case of asymmetrical molecules the work of adhesion towards water is greater as compared with the work of adhesion when symmetrical molecules are involved. This excess work of adhesion of asymmetrical molecules towards water may find its explanation in the rupture of very strong electromagnetic fields, while in the case of separating water molecules from water only the weakest fields are ruptured. It is evident from the above that spreading of liquids in the process of foam formation becomes an important factor, especially when solid agents participate in the production of a foam system where forces of adhesion and cohesion come into play.

### 5. Conclusions

1. Capillary-active substances promote foaming at a definite concentration.
2. An abrupt drop between the concentration of the boundary layer and the solute is a requisite for foaming.
3. The relation between the solubility and the surface activity of a foamer determines whether it is strong or weak.
4. Foaming, as a rule, occurs only in the case of a heterogeneous Gibbs layer.
5. Colloidal solubility does not favor foaming. The formation of large molecular aggregates interferes with the principle of heterogeneity.
6. The Gibbs layer in foam systems may differ with respect to the establishment of a state of equilibrium. A distinction is made between thermodynamic and chemical irreversibilities.
7. The wetting and spreading factors are especially emphasized in case the foamer is a solid. Substances less wetted by water adsorb capillary-active substances from aqueous solutions to a greater degree.

### III. THE STABILITY OF FOAM SYSTEMS

Although foam systems are closely related in their properties to emulsions, yet in their stability they differ greatly. The life of foams is short in comparison with the life of emulsions. Foams are rather labile formations because in other dispersion systems a true static equilibrium is possible,—the kinetic energy of molecules creates a considerable opposing action. This kinetic resistance is absent in foam systems.

On the other hand, combined with a much greater surface increase, foam formation greatly opposes the tendency of the system to reach a stable state

TABLE 1  
*Stability of foams according to observations of Plateau*

SUBSTANCE	STABILITY	EXPERIMENTAL CONDITIONS
<i>Class I:</i>		
Distilled water.....	1- 7 seconds	Atmosphere saturated with vapor
	1- 12 seconds	Atmosphere dried with sulfuric acid
Glycerol.....	- 2 seconds	
Sulfuric acid.....	1- 28 seconds	Atmosphere dry
Nitric acid.....	- 1 second	Atmosphere saturated with vapor
Ammonium hydroxide.....		Atmosphere saturated with vapor
Saturated solution of tartaric acid.....	1-142 seconds	
Potassium nitrate.....	- 6 seconds	Atmosphere saturated with water vapor
Sodium carbonate.....	- 26 seconds	Atmosphere saturated with water vapor
	1- 30 seconds	Atmosphere dried with calcium chloride
<i>Class II:</i>		
Olive oil.....	0- 7 seconds	
Lactic acid.....	1- 18 seconds	
Acetic acid.....	0- 8 seconds	Atmosphere saturated with vapor
Turpentine.....	1- 6 seconds	
Benzene.....	Maximum -12 seconds	
<i>Class III:</i>		
Marseille soap.....	30 seconds-2 hours	Atmosphere saturated
Domestic soap.....	5- 14 seconds; colorless phase of water vapor; stable about 90 minutes	
Potassium colophane solution.....	10- 30 minutes	
Saponin solution.....	25- 40 minutes; colorless phase; maximum stability, 12 hours	
Albumin solution.....	Many hours; colorless phase; stability may be several days	
Ferric acetate.....	15- 30 seconds; colorless phase; maximum stability, about 24 hours	

TABLE 2  
Maximum foam duration for various substances

AQUEOUS SOLUTIONS OF	OPTIMUM CONCENTRATION	MAXIMUM FOAM DURATION
	<i>moles per liter</i>	<i>seconds</i>
Propionic acid.....	0.26	11.0
Lactic acid.....	1.00	7.5
Propyl alcohol.....	0.34	11.0
Glycol.....	6.00	1.5
Glycerol.....	6.00	3.0
Cane sugar.....	0.50	0.5
Ethyl alcohol.....	0.28	5.0
Propyl alcohol.....	0.34	11.0
Isobutyl alcohol.....	0.09	12.0
Isoamyl alcohol.....	0.036	17.0
<i>tertiary</i> -Amyl alcohol.....	0.034	10.0
Heptyl alcohol.....	0.0007	8.0
Octyl alcohol.....	0.0003	5.0
Formic acid.....	0.45	4.0
Acetic acid.....	0.20	8.0
Propionic acid.....	0.25	11.0
Butyric acid.....	1.00	18.0
Valeric acid.....	0.015	9.0
Caproic acid.....	0.0075	13.0
Heptylic acid.....	0.0015	16.0
Caprylic acid.....	0.00025	12.0
Nonylic acid.....	0.00007	5.0
Ethylamine.....	0.40	12.0
Aniline.....	0.10	11.0
<i>p</i> -Toluidine.....	0.04	6.0
Phenol.....	0.10	12.0
Benzyl alcohol.....	0.10	10.0
<i>m</i> -Cresol.....	0.025	9.0
Nitrobenzene.....	0.005	6.0
Benzene.....	—	0.2
Acetaldoxime.....	0.37	10.0
Paraldehyde.....	0.03	9.0
Acetone.....	0.50	2.5
Methyl propyl ketone.....	0.05	3.0
Ethyl propionate.....	0.01	2.5

by diminution of the surface with a resultant decrease in surface tension. But this surface tension decrease is subjected to lasting changes in time, and this is the cause of having a dynamic, not a static, equilibrium.

The work of Plateau contains experimental data with respect to the duration of a foam from various foaming solutions; this is summarized in table 1.

Recently the relation between foam duration and surface activity was studied; the results showed that an increase in foam duration is parallel with an increase in surface activity (3). This was the case with fatty acids up to butyric and with alcohols up to isoamyl. Maximum foam duration

TABLE 3  
*Maximum foam duration for various substances*

AQUEOUS SOLUTIONS OF	MAXIMUM FOAM DURATION	CONCENTRA- TION AT MAXIMUM FOAM DURATION	SURFACE TENSION OF SOLUTIONS AT MAXIMUM FOAM DURATION
	<i>seconds</i>	<i>moles per liter</i>	<i>dynes</i>
Ethyl alcohol.....	5	0.28	66
Propyl alcohol.....	11	0.32	56
Isobutyl alcohol.....	12	0.09	56
Isoamyl alcohol.....	17	0.036	54
<i>tertiary</i> -Amyl alcohol.....	10	0.034	61
Heptyl alcohol.....	8	0.0007	68
Octyl alcohol.....	5	0.0003	64
Formic acid.....	4	0.45	70
Acetic acid.....	8	0.20	69
Propionic acid.....	11	0.25	61
Butyric acid.....	18	1.0	33
Valeric acid.....	9	0.015	64
Caproic acid.....	13	0.0075	57
Heptylic acid.....	16	0.0015	60
Caprylic acid.....	12	0.00025	66
Nonylic acid.....	5	0.00007	70
Benzyl alcohol.....	10	0.1	59
<i>m</i> -Cresol.....	9	0.025	62
Aniline.....	11	0.1	62
<i>p</i> -Toluidine.....	6	0.04	63

for fatty acids and alcohols is presented in tables 2 and 3 (experiments by Bartsch). The surface tension of aqueous solutions compared by Bartsch, using a concentration corresponding to the maximum foaming ability as a criterion, shows great deviations accounted for by isocapillarity of the surface-active substances.

#### A. FACTORS IN THE STABILIZATION OF FOAM SYSTEMS

With regard to factors influencing the permanency of a foam, although the degree of stability is less, a similarity to emulsions is evident.



The following factors imply the conditions under which an emulsion system becomes stable, namely, low interfacial tension, optimum concentration, viscosity, behavior with electrolytes, action of protective colloids, presence of a foam agent, existence of a Gibbs adsorption layer and its thickness, and lastly orientation of molecules in the foam interface. There is also a certain peculiarity in the behavior of foam systems, as compared with emulsions, which may be attributed to their difference in character. As mentioned above, from a purely thermodynamic standpoint it follows that foams cannot be as stable as emulsions. Electrical forces contributing to the stability of emulsions are not so important in the case of foams where capillary forces apparently play a greater part. The dispersion factor does not act in the same direction in the case of foams as in emulsions.

Foam systems may be differentiated as follows (3):

1. Those originating by shaking water and surface-active organic substances,—two-phase foam systems.
2. Those obtained by shaking emulsions of organic substances, as well as lyophilic foams,—three-phase foam systems.
3. Those systems in which the foam agent is a colloiddally dissolved phase and acts as an integral part of the dispersion medium in the foam system so that it can not be distinguished from two-phase foams in which the molecularly dissolved foaming agent is also an integral part of the dispersion medium.
4. Those in which a solid phase is present occurring as an independent component of the system and in macroscopic quantities similar to the gas phase.

To comprehend the measurable stability of foam systems, first those systems must be considered in which the foam agent is present as an electrolyte, molecularly or colloiddally dispersed, as a protective colloid, or as a solid foam former. In this regard a certain analogy may be drawn between foams and emulsions. Permanent foams, as well as permanent emulsions, under ordinary conditions of temperature and pressure are obtained by using foam and emulsifying agents. Further, the degree of stability of a foam depends upon the nature of the components participating in the system and whether or not foam originates by shaking solutions of lyophilic or lyophobic substances. While the first foam systems are more stable and less sensitive to added electrolytes (analogous to lyophilic sols), the lyophobic substances render the foams unstable and are very sensitive to electrolytes.

### *1. The influence of electrolytes*

Lyophilic foams have more outstanding properties, such as viscosity and elasticity, and the influence of electrolytes is so small that in most cases it

cannot be proved. This phenomenon is explained by Bartsch by the fact that the foam walls of these systems consist not of the sol, but of the coagulated sol or gel of the lyophilic colloid, as shown by Ramsden and others. The dispersion medium is in a rigid or semi-rigid state, which acts opposite to the change in dispersion so that no precipitation of electrolytes may be expected. Such foam systems may be sensitized by means of surface-active substances. Bartsch proved experimentally that foam formed by shaking a 0.001 per cent saponin solution, after the addition of 0.5 per cent of isoamyl alcohol, was found to be sensitive to electrolytes. The behavior of unstable foams with respect to electrolytes is similar to lyophobic sols. The only difference is the time of stability; lyophobic sols are stable when compared to lyophobic foams. The speed of coagulation caused by mechanical changes occurring in foam walls is frequently so high that it exceeds the speed of precipitation of electrolytes. Some of these foams are sufficiently permanent and undergo only coagulation induced by electrolytes. Foam duration is usually expressed in the number of seconds from the moment of interruption of shaking of the foam-forming liquid to the complete disappearance of the foam formed. Bartsch investigated foam systems of amyl alcohol, quinoline, turpentine, and eucalyptus oil. The similarity between lyophobic sols and lyophobic foams is obvious from an investigation of the hydroxyl ion influence upon the stability of foams. The hydroxyl ions were found to produce a considerable increase in the stability of foam systems, and a decrease in the number of foam bubbles, increasing the degree of dispersion of the gas phase. On the other hand, hydroxyl ions had a peptizing effect upon the foam system, but this action was limited by the degree of concentration. Bartsch's investigation proved that the duration of a foam formed from an isoamyl alcohol solution was decreased from 18 to 15 seconds by the addition of the following concentrations of electrolytes:

<i>Electrolytes</i>	<i>Millimoles per liter</i>
NaOH.....	30.0
NaCl.....	1.0
Na <sub>2</sub> SO <sub>4</sub> /2.....	0.8
H <sub>2</sub> SO <sub>4</sub> /2.....	0.04
BaCl <sub>2</sub> .....	0.05
AlCl <sub>3</sub> .....	0.0005

The stability threshold displacement for various electrolytes is well pronounced.

Potassium carbonate and potassium citrate were found to be excellent peptizers. Although the peptizing action is also a function of concentration, Whitney and Straw (71) found that for the peptization of emulsions the favorable concentration lies between 0.2 and 80 millimoles per liter; foam systems show an optimum action at concentrations of 0.1 to 10 millimoles per liter.

## *2. The action of protective colloids*

Further analogy between lyophobic foams and lyophobic sols is found in their behavior with respect to the action of protective colloids. The influence of protective colloids upon the stability of foam systems has also been discussed and proved experimentally by Bartsch. Lyophobic foam systems may be protected from the precipitating action of electrolytes by protective colloids. The actual concentration of the protective colloid in a foam system is doubtful and differs from the starting solution. On account of the decreasing surface tension, lyophilic colloids gather at the boundary surface of the solution. The presence of these substances in a foam is shown by an increase in its stability. Approximately equal amounts of various protective colloids are necessary to increase foam duration to the same extent. Bartsch found that the smaller the concentration, the more surface-active is the protective colloid. Very strong surface-active substances show an extreme deviation (see table 12 in the paper by Bartsch (3)). The relative protective number for foams with respect to the sequence of the protective action was found to be in complete agreement with that of gold numbers for equal protective colloids. A protective colloid may act on a foam system in a direction opposite to the one mentioned above. Since precipitation in the interface is considered a fundamental condition for the formation of a three-phase foam system, it seems reasonable to assume that those protective colloids which prevent this precipitation make the formation of three-phase foam systems impossible. The protective colloids at certain concentrations displace completely the solid phase from the foam system, and in this case an unstable two-phase foam system is formed instead of the three-phase. The alternative explanation favored by Bartsch is that a stable two-phase system develops when the protective colloid is present in the solution in a high enough concentration, because protective colloids themselves are able to build up a stable two-phase foam system. Flotation practice makes use of this phenomenon when the behavior of various ores against certain protective colloids is determined by measuring the weighed amounts of the corresponding ores which may still be adsorbed by the foam at equal concentrations of the colloid.

## *3. Viscosity as a factor in the stabilization of foam systems*

Bartsch also worked out the idea of increasing the stability of foams by the addition of substances increasing the viscosity. He carried out experiments with isoamyl alcohol solutions to which various amounts of glycol and glycerol were added. The result was that only very high concentrations of about 60 per cent increased the stability markedly. This

was explained by the action of the "removal law" of Michaelis and Rona (39), according to which glycerol or glycol was partially removed from the foam walls by the presence of surface-active substances. A soap-saponin solution may serve as an example of the "law of removal." In this case, the soap solution has a smaller surface tension than the saponin solution, and this results in the removal of soap from the interface by saponin. The fact that a saponin solution loses its ability to form a stable foam on the addition of alcohol may be partially explained by the same removal process but, on the other hand, it is essential that saponin be really soluble in alcoholic solutions and not form skins.

Quincke (47) was the first to describe the peculiar phenomenon that ethyl ether, present even in small amounts in the surrounding atmosphere, destroys the foam of a good foaming liquid. Again the capillary-active ether removes the foam agent in the liquid. The skin-forming abnormally adsorbed substances, such as soap, saponin, lyophilic foam agents (albumin and gelatin), the adsorption of which is not influenced by capillary-active substances taken in minute amounts, may increase considerably the stability of isoamyl alcohol foam. An amount of 0.000001 per cent of soap increases the foam stability twofold, while 0.001 per cent of gelatin is required for the same action. This depends on the surface activity of the corresponding substances. Stability also depends to a certain extent on whether the foam contains large or small bubbles, and whether the bubbles are elastic or rigid. Certain substances added to a foam-forming solution not only increase the stability by an increase in viscosity, but render a certain elasticity by softening the interface skins, thereby preventing the coalescence of bubbles in a foam.

Boys (6) reports an experiment in which one drop of a soap solution added to 30 g. of a saponin solution was sufficient to prevent the rigidity of saponin bubbles. On the addition of three drops, the surface skin became movable and elastic, as in the case of pure soap solutions. Glycerol added to a soap solution induces foaming, increases viscosity, and, at the same time, assists in the production of soft and flexible skins. Therefore saponin solutions with glycerol resemble, in their behavior, pure soap solutions. Elastic foams and those with large bubbles appear to be more stable. It is probable that an explanation for the stability of durable foams as, for example, those produced by the sea, by beer, or by aqueous solutions of saponin and soap, must be sought partially in the formation of very viscous or semi-rigid or gel-like membranes at the interface. Wheat gum used in the beer industry, as well as other dextrin-like substances, acts similarly to glycerol, increasing the viscosity and producing softer and more elastic skins. That the various degrees of stability of a foam often depend on the properties of the bubbles forming the foam has been proved.

Whether the bubbles are elastic or rigid depends chiefly on the nature of the films. Soap solutions are distinguished by the formation of very elastic bubbles, which show a smooth flexible surface. Contrary to this, bubbles of saponin solutions are much more rigid; when a bubble collapses, a pleated sac is observed. The difference between the elastic lamina of soap and the rigid type of saponin becomes evident from the rearrangement in colors which takes place with a continuous change in the thickness of the lamina. In saponin foams, the laminae are so rigid that certain color spots are retained unaltered upon a given lamina. In the tendency towards equilibrium, some laminae burst or unite to form new ones. But still the history of the laminae in question may be read from these invariable color spots.

#### *4. The action of finely dispersed solids*

While the behavior of gas-liquid and liquid-liquid systems in themselves is evidently different, nevertheless certain analogies are met with. Both emulsion and foam systems may be stabilized, not only with agents which are molecularly or colloiddally dissolved, but with finely dispersed solids, first applied by Pickering for emulsion systems. The latter may replace solutions in the interface and build up a boundary layer to increase the stability of these systems. Donnan (11), speaking of the stability of foam, states: "If air is violently churned up with water, only comparatively large bubbles are produced and these quickly rise to the surface and burst. If a very small amount of a substance which concentrates at the air-water interface is added, an almost milk-white air emulsion of small bubbles is produced which rise to the surface and produce a relatively durable froth." According to Lord Rayleigh (51), who discussed this phenomenon, it is clear that diminution of interfacial tension facilitates the subdivision or the dispersal of air. The existence of the surface layer confers a certain amount of stability on the resultant foam, since it will give rise to forces which resist the thinning of a bubble wall. Any sudden increase in the surface will produce a momentary diminution in the concentration or thickness of the interface layer and hence an increase in surface tension, which will persist until the normal thickness or concentration is readjusted by diffusion of molecules from the inside volume,—a process which in very dilute solutions occupies a perceptible time. That this explanation of Lord Rayleigh is correct, says Donnan, may be seen from the fact that very often stronger solutions of the same surface-active substance would not foam at all. In this case the readjustment of the equilibrium thickness or concentration in the gas-liquid interface occurs with such rapidity that practically no rise in surface tension and hence no counteracting force comes into play (24). On the other hand, when higher concentrations are

taken, aggregates of gas bubbles and solid particles can no longer be formed by shaking, because the interfaces formed thereby are already occupied by molecules of the surface-active substance.

Dynamic equilibrium (2) established between the speed of formation of the adsorbed layers of the foam-forming agent or stabilizer at the solid-liquid boundary on one side and the speed of such layer formation at the gas-liquid interface on the other side determines to a certain extent the stability of this three-phase foam system in the case of higher concentrations of the foam agent. As has been demonstrated in the experiments of Hoffmann (25) and others, small solid particles may contribute to the stabilization of a froth or foam as in the case of the mineralized frothers or the ore flotation process by the preferential aggregation in the interface between two phases. Bechold, Dede, and Reinders (4) state that solids prevent coalescence of bubbles in foams, as well as in emulsions. Solids prevent coalescence of droplets, thus increasing the stability of a system. The degree of dispersion of the solid was found to be of even greater significance in foam systems than in emulsions. Substances in a very fine degree of dispersion increase the foam stability much more than substances of a coarse grain dispersion. There is an optimum dispersion of the solid phase for the stabilization action between colloidal and coarse distribution (2). For example, when lead glance which has been passed through a 900-2500 mesh sieve is used as a foam agent, it increases foam stability to 60 seconds; when lead glance which has been passed through a 9000-12,000 mesh sieve is used as a foam agent, the life of the foam formed is increased to several hours. Further, the amount of the stabilizer, whether it is a liquid acting as a foam-former or a solid, is decisive for the formation of a stable foam. The magnitude and stability of aggregates of gas bubbles and solid particles are determined by the concentration of the foam agent, and the nature and affinity of the foamer and the solvent for the solid phase.

## B. FORCES INVOLVED IN MAINTAINING STABILITY

### 1. *Electrostatic and capillary forces*

Among the forces involved in maintaining the stability of a foam system are electrostatic and capillary forces. Concerning the nature of forces producing an attraction between solid particles and a gas, it is known that all components of a three-phase foam system are charged with electricity of the same sign, gas bubbles being charged negatively against water, therefore stability depends on the valence and the adsorption ability of the cations. The electric influences are indirect rather than direct and seem to be responsible only as to which boundary phase the particles may enter and the degree to which they may coagulate and change the structure of

the foam. The charge of the gas phase may be determined by the indirect method, using the ion precipitation series of colloid chemistry. Cations produce a precipitation in various foam systems, according to their valency, and these systems have been found to be negatively charged. Bartsch determined the charge on a gas bubble in a foam system, but the potential drop at different boundaries in the system has not been proved. It is still undecided whether the electrokinetic potentials or the phase boundary forces are dealt with. Where gases are acting, the balloelectric effect should be considered. The fuzzing effect of Lenard (15, 33) is based upon a similar phenomenon fundamental to balloelectricity. A balloelectric effect occurs when gas bubbles pass through a liquid and then leave the surface. The electrical energy from the double layer reaches the gas phase and gives to it a certain charge. Balloelectricity implies streaming potentials, but in this case the charges at the boundary of the liquid (gas) move readily; therefore, not only a displacement, as in the case of streaming potentials, but rather a destruction of the double layer by the tangentially-acting mechanical forces takes place and the electricity carried by both layers (external and internal) reaches the gas. The gas phase was found to be negative, while electrolytes, acids, and certain active substances (8) produce a change in the charge even for small concentrations. In general, electric influences in foam systems have not as yet been thoroughly investigated.

Taggart (62) studied the changes of the potential under the influence of electrolytes. He permitted a small gas bubble to migrate cataphoretically around the longitudinal axis of a rotating glass cylinder between two electrodes and determined the direction of migration and its speed under various conditions. He found that the gas bubble in pure water was charged negatively. The velocity of air, oxygen, and hydrogen spheres in water was found to be about  $4 \times 10^{-4}$  cm. per second per volt per centimeter. The presence of hydrogen ions decreased the boundary phase potential to zero without a change in the sign of the charge. Tri- and tetra-valent cations produced a change in the sign of the charge. A sphere of gas in a solution was found to change the sign of its charge as it diminished in size by being adsorbed into the solution. Surface-active substances decreased the cataphoretic speed according to Traube's rule. In general, the interfacial potential differences are of great importance and play a fundamental part in determining the stability or instability of the colloidal state of matter. In the case of foam interfaces, potential differences are not perhaps of such great importance, but since very little work has been carried out in this direction, our opinions as to the importance of this factor must be held in abeyance.

It has been shown only in the case of an air-water surface that there

exists an electrical separation or potential difference in the interface layer and that certain substances can produce pronounced variations or even reversal in sign for this electrical double layer. There may be a possibility for direct measurement of the potential drop between single phases of a foam system using the Haber-Klementschevitch glass electrode. The influence of charge upon attraction between components of a foam system is marked only in case other forces, such as capillary forces, are weak. The fact that gas bubbles adhere to solid particles in spite of electrostatic repulsion of the components indicates that the more powerful capillary forces participate to a great extent in maintaining permanency.

The coalescence of two drops takes place at the boundary partition, where a diminution of surface occurs and where capillary forces originate tending to tear apart these boundary sections and to counteract their coalescence. Capillary forces act against molecular attraction,—a cohesion between molecules when dealing with the same kind of molecules and an adhesion when different ones are involved. These forces existing between molecules of similar and different natures, more than the electrical elementary forces, must be considered in the case of three-phase foam systems.

Talmud and Suchowolskaja (63), in their recent study of elementary foams, pointed out that stability depends rather on the degree of hydration and that the slipping velocity of the water layers in the walls of laminae is a determining factor of stability because substances stabilize before the adsorption layer is obtained.

## *2. The adsorption layer in the interface of a foam system and its thickness as a factor influencing stability: the Langmuir-Harkins orientation theory*

As in the case of emulsion systems, so in three-phase foam systems,—for example, amyl alcohol, solid agent (foamer), oleic acid (stabilizer)—a relationship appears to exist between stability, the presence of an interface layer, and its thickness.

Adsorbed molecules are held in direct contact with both gas and water in a foam system. Some adsorption surfaces are extremely viscous or even rigid (saponin and nearly all proteins), and the adsorbed particles are in mutual contact. Others are freely mobile (quinine and soaps), and the particles are either more widely spaced or so oriented that only cohesion between a liquid comes into play. In many cases where adequately permanent bubble films are obtainable, the interference colors developed as the film becomes thinner are disposed in horizontal bands when the surfaces are mobile, but chaotically disposed when the surfaces are rigid.

It is known that gases or vapors can be condensed or adsorbed by solid and liquid surfaces. The question arises whether the formation of primary



unimolecular films ever occurs in such cases. Langmuir (30) measured the adsorption of a number of gases at low temperatures and pressures on defined surfaces of mica and glass and concluded that the maximum quantities adsorbed are always somewhat less than the amounts to be expected in a unimolecular surface layer. Carver (7) found similar results measuring the adsorption of toluene vapor on known glass surfaces.

This view that the maximum adsorption from the gas phase cannot exceed a unimolecular layer has, however, been much criticized. Thus Evans and George (13), on the basis of their own measurements of the adsorption of gases on a known surface of glass wool combined with data obtained by Mulfarth (42), concluded that the adsorption layer may be several molecules thick. Reference may also be made to an article by Wilson (72), who deduced from the results of the adsorption of oxygen and nitrogen at low temperatures that these gases are held in layers several molecules thick.

Another type of surface layer formation is at the surface of liquids where a substance dissolved in a liquid concentrates preferably at the liquid-air or liquid-vapor interface. Gibbs (16), and later Thomson (64), showed that if a dissolved substance (in relatively dilute solution) lowers surface tension, it will concentrate at the surface. That such a phenomenon actually occurs in foam has been demonstrated qualitatively in the experiments of Hall (18), Zawidski (73), and Kendrick and Benson. If various experimental values are considered, then evidence is produced that they are not materially different from those found by Langmuir and Adams for the oriented unimolecular layer of practically insoluble fatty acids resting on the surface of the water. Harkins' theory, regarding monomolecular films produced on water, points out that these films occur only when the spreading coefficient has a relatively high value. Since these high values result only when the spreading substance contains a polar group in its molecule, the presence of such a group is essential to a monomolecular film for spreading on water, but not at all essential for the production of a thicker film. Donnan explained the fact that the thickness values in some cases are greater by the partially or completely vaporized state of the adsorption layers.

Adams and Marcelin (1) made the important discovery that the unimolecular surface films which they investigated may pass rapidly, on increase in temperature, from the state of solid or liquid surface films to the state of vaporized surface films in which the "juxtaposed" molecules become detached from each other and move about with a Brownian or quasimolecular motion. These investigators attributed this behavior to the thermal agitation of the water molecules to which they were attached.

Bartsch first emphasized the fact that the stability of three-phase foam systems, analogous to three-phase emulsion systems, depends upon the dimensions of the adsorbed Gibbs layer of the stabilizer. He calculated the thickness of the adsorbed layer from data obtained for the weight of the solid phase, the specific gravity of the solid phase, and the size of the particle given by the degree of dispersion, and found that the adsorbed layer under stable conditions may run up to five to ten molecules thick with a layer of galena, two molecules with copper pyrites, and five to seven molecules with zinc blende.

This calculation does not consider the fact that the adsorbed surface is actually much larger than the calculated surface of particles due to possible rupture of the adsorbed surface. Also, a quantitative adsorption has been assumed, while from these determinations it follows that in very dilute oleic acid solutions the adsorption was relatively less than in concentrated solutions. Adsorption of oleic acid at the boundary surfaces of air bubbles was not accounted for. The stability of three-phase foam systems is a function of the foaming agent or stabilizer, that is, the power of adhesion between solid particles and air bubbles is determined by the adsorption layer. In order to render three-phase foam systems stable, the thickness of the solid layer or the oleic acid concentration should be such that the surface of the solid phase is covered with an oleic acid layer; this surface layer has none of the properties of oleic acid as a whole. Since the thickness of the layer considered was less than four molecules in diameter, the probability of a monomolecular layer was assumed. Beyond the optimum thickness, the bond between solid particles and air bubbles is weakened and the stability of the system decreased. Freundlich (15), discussing the action of solids upon foams, explains the resistance ability of the foam structure against solids as follows: When a solid is wetted by such solutions as soaps and saponin, it is "sucked" into the edge canals which are rich in liquid without destroying the structure of the foam externally.

Bartsch has defined conditions for the stability of a foam, taking into consideration the orientation theory of Langmuir (31) and Harkins (21).

The orientation theory postulates that molecules in the interface layer assume a definite direction. The atomic groups (carboxyl and oxyamido) of a molecule rich in residual valences are so placed that an almost complete saturation occurs in the boundary. When the boundary is not saturated by any residual valences, then the molecule directs the most saturated atomic group to this boundary, that is, the carboxyl group of oleic acid is directed to the solid phase, while the methyl group is directed towards the outside. When a methyl group comes in contact with a water-air boundary, it displaces the water molecules from the boundary layer because the

interfacial surface tension in the presence of the methyl group is decreased, i.e., the water molecules are subjected to the expelling influence of the carboxyl group when competing for the surface of the solid phase.

The energy of the interface has a tendency to reach the lowest possible value, and this is the force which causes cohesion of solid particles to air bubbles. When only one molecular layer of oleic acid is present between air bubbles and solid particles, the intensity of adhesion is produced by the stronger forces existing between the residual valences of the solid phase surface and the forces of the carboxyl groups of the oleic acid. The latter forces are essentially greater than the forces of cohesion acting between the oleic acid molecules, as shown by Harkins. In the case of octane, the entrance of an oxygen atom into the molecule increases its cohesion up to 2 per cent, while the adhesion to water increases the cohesion up to 65 per cent. But the oleic acid is adsorbed not only by the solid phase, but also by the water-air boundary, depending upon which of the boundaries the oleic acid molecules have been previously adsorbed or, in other words, upon the rate of adsorption for the two boundaries. In the case of high oleic acid concentrations which are sufficient to surround gas bubbles as well as solid particles with an oleic acid layer, a certain attraction is always observed. Bartsch explained this as follows: In shaking an oleic acid emulsion with air in the presence of solid particles, the water-air boundary is always renewed. The rate of adsorption of the water-air interface and the rate with which an oleic acid droplet spreads upon the surface of a gas bubble is less than that of the adsorption by the solid phase.

Langmuir's conception regarding spreading, in this case, is based on the attraction existing between carboxyl groups and the molecules of the surface of the water or the surface of the solid phase. This attraction for the surface of a solid phase is greater than that for the surface of water, because the water molecules are displaced by solid particles wetted by oleic acid. The difference between the rate of adsorption and the rate of spreading of the oleic acid at the boundaries is due to the fact that even at higher concentrations of oleic acid the water-air boundary is not completely covered with a closed oleic acid layer, so that there will always exist a possibility for the formation of gas bubble complexes and solid particles. A gas, a liquid, and solid particles, held together by oriented oleic acid layers, comprise stable foams.

A monomolecular layer of oriented oleic acid molecules binds the aqueous solution to the gas and the solid surface, as well as the gas to the solid surface. The relation between the thickness of the interface layer and the stability of a foam, according to the orientation theory, may be described thus: As a result of the tendency of the free energy of the interface to acquire the lowest possible value in the case of low concentrations, the

intensity of adhesion between solid particles and gas bubbles decreases with a decrease in the number of molecules of the foamer or stabilizer inducing adhesion. The intensity of adhesion between solid particles and gas bubbles is due to saturation of groups of molecules of the stabilizer rich in residual valences. These valences have their origin in the boundaries of the solid phase, as well as in a decrease in surface tension produced by saturated groups of substances in the gas-liquid boundaries.

In the case of high concentrations exceeding the optimum conditions of a state of permanency, the intensity of adhesion is the resultant of cohesion existing between molecules of the foamer in the adsorbed layer which, in this case, is made up of a considerable number of layers of molecules of the adsorbed substance.

The influence of cohesion becomes more marked the greater the number of layers of molecules. Orientation requires unsymmetrical molecules; for the most stable form of a system, it is necessary that it give the lowest free energy. Therefore the molecules orient themselves in the interface so that the paraffin hydrocarbon chain remains in the oil, while the active polar group (COOH), which exercises the greatest electromagnetic force, is in the water.

At the surface of the solid phase the molecules of water are subjected to the displacing influence of carboxyl groups; at the water-air boundary, the displacing influence of methyl groups towards oleic acid is effected.

### C. CONCLUSIONS

1. Foams are labile formations with kinetic resistance lacking.
2. A dynamic, not a static, equilibrium is characteristic of foam systems. A great surface increase in foam formation counteracts stability with a decrease in surface tension.
3. The nature of the components participating in a system determines foam stability; lyophilic foams are more stable than lyophobic.
4. The rate of adsorption in the boundary accounts for the stability of foams.
5. Some ions may produce a considerable increase in the stability of foam systems, increasing the dispersion of the gas phase, while other ions act as peptizing agents.
6. For solid foamers, foam stability increases with the degree of dispersion.
7. High concentrations of substances added to increase the viscosity of a system increase its permanency.
8. Certain substances possess the ability to displace foamers from the interface.
9. Certain substances, such as glycerol, have a greater influence in the

production of soft and flexible skins than on viscosity. Elastic foams composed of large-sized bubbles are considered more stable than those of small-sized bubbles; this is not in alignment with other colloidal systems.

10. Electrostatic and capillary forces are involved in maintaining the stability of a foam.

11. The relation between the thickness of the interfacial layer and stability is analogous to that of emulsion systems. The Langmuir-Harkins orientation theory is applicable to foam systems.

#### IV. TRANSITION TO THE UNSTABLE STATE

##### A. FORCES ACTING IN THE THINNING OF A FILM

The interior of films composing foams is subjected to a continuous diminution in thickness by internal currents due to gravity and suction at the edges. The weight of the lamina has the tendency to make it thinner by forcing the molecules to slip continuously from the top to the bottom of the bubble. On the other hand, the liquid between two surface skins of a lamina moves upwards very slowly following the force of gravitation. The friction is great because, according to Poiseuille's law, the speed with which a liquid moves in a capillary is proportional to the fourth power of the radius and is, therefore, extremely small in such a small cross section.

Liquids behave differently. Gibbs (16) assumed that the motion of the liquid through which a bubble of air is rising is similar to that of a liquid running out from between two surfaces. This motion of the liquid may be greatly retarded by its viscosity as soon as the space between the two surfaces is reduced to the thickness of a lamina. The effect of an increase of the surfaces in increasing their tension becomes greater and more permanent as the quantity of liquid available for supplying the substances forming the increased surfaces diminishes. Gibbs gives a rough estimate of the amount of motion possible in the interior of a liquid lamina with respect to its exterior by calculating the descent of water between parallel vertical planes when the motion of water is reduced to zero. He uses the coefficient of viscosity determined by Helmholtz and Piotrowsky (22) and obtains  $V = 581D^2$  where  $V$  denotes the mean velocity of water expressed in millimeters per second and  $D$  denotes the distance in millimeters between the fixed planes, which is supposed to be very small in proportion to their other dimensions. For the same temperature, the descent of water in long capillary tubes is  $V = 337D^2$  and for descent between parallel planes it is equivalent to  $V = 899D^2$ . Gibbs draws the conclusion that in a film of a liquid which has a viscosity and a specific gravity not very different from that of water, the mean velocity of the interior relative to the surfaces will probably not exceed  $1000D^2$ . This is a velocity of 1 mm. per second

for a thickness of 0.1 mm., 0.6 mm. per minute for a thickness of 0.001 mm., corresponding to the interference light of red of the fifth order in a film of water, and 0.036 mm. per hour for a thickness of 0.0001 mm., corresponding to white of the first order. Such a small internal current is evidently consistent with great persistence of the film, especially in those cases where the film can exist in a state of the greatest tenuity. On the other hand, Gibbs states that the above equations give so large a value of  $V$  for a thickness of 1 mm. or 0.1 mm., that a film evidently can be formed without drawing up any great weight of liquid, and any such thicknesses as these can have only a momentary existence.

Besides the action of gravity on the interior of a film, there is another cause producing changes, namely, the edge where the film terminates. At such an edge, there is generally present a liquid mass continuous in phase with the interior of the film bounded by concave surfaces. The pressure at the edge is, therefore, less than that in the interior of the lamina. Thus the liquid exerts a strong suction upon the interior of the lamina by which its thickness is rapidly reduced. Unless the lamina is very viscous, its decrease in thickness near the edge causes a rapid upward current on each side, while the central portion slowly descends. These processes rapidly reduce the thickness of the film. The volatility of the principal component is another example of the downward tendency of the interior of the film.

If in the case of a film of soap and water the humidity of the surrounding atmosphere is such that the vapor pressure of water at a level between the top and the bottom of the film has the same value in the atmosphere as in the film, then evaporation takes place in the upper portions and condensation in the lower portions. These processes occasion currents of diffusion, the effect of which is to carry the moisture downward. In the event the vapor pressure of water in the atmosphere differs little from that in the film, the processes described would not be sufficiently rapid to be of importance. But in the case of a film of soap and water in a dry atmosphere, or a film of soap and water with glycerol in a moist atmosphere, the effect of evaporation or condensation is not to be slighted. In the first case, a decrease in the thickness of the film is accelerated; in the second, it is retarded. In the case of a film containing glycerol, it is to be expected that the condensed water will not replace the fluid carried down by the internal current, but that a combination of these two processes will tend to remove the glycerol from the film. When a component which greatly diminishes the tension of a film is volatile, the effect of evaporation and condensation may be considerable, even when the mean value of the vapor pressure for that component is the same in the film as in the surrounding atmosphere. Gibbs illustrates this by using two components,  $S_1$  and  $S_2$ , and states that the

vapor pressure in a film must vary with the weight  $Z$ , since surface tension varies, and these variations very nearly satisfy the equation:

$$\gamma_1 \frac{d\mu}{dZ} + \gamma_2 \frac{d\mu_2}{dZ} = 0$$

$\gamma_1$  and  $\gamma_2$  denoting the densities of the components  $S_1$  and  $S_2$  in the interior of a film. The variation of the potential of  $S_2$  as we pass from one level to another is, therefore, as much more rapid than that of  $S_1$  as its density in the interior of the film is less. If the resistances restraining evaporation, transmission through the atmosphere, and condensation of the two substances are the same, then these processes will proceed more rapidly with respect to  $S_2$ , and the values of  $\frac{d\mu}{dZ}$  and  $\frac{d\mu_2}{dZ}$  will have opposite signs, the tendency of  $S_1$  being to pass downward and that of  $S_2$  to pass upward. Moreover, the evaporation or condensation of  $S_2$  will produce a much greater effect than evaporation or condensation of the same quantity of  $S_1$ . If condensation of  $S_2$  takes place at the top of the film, it will cause a decrease in surface tension and an expansion in that part of the film by which its thickness is reduced as in evaporation of  $S_1$  (16).

#### B. FACTORS INVOLVED IN THE TRANSITION TO THE UNSTABLE STATE

Thus the general conditions of the persistence of a liquid lamina are such that the substance which causes a decrease in surface tension in the lower parts of the film must not be volatile. Volatility influences rupture, for it is understood that when superficial molecules are lost by evaporation, their intrinsic cohesion must have been very small. In spite of the mobility of the superficial layers, little force is necessary to produce rupture; motion (evaporation) in the external molecules favors rupture. Plateau, in his theory on foams, assumed that in order for a liquid to be able to form bubbles, develop a large film and yet be resistant, it is necessary that the viscosity of the superficial layers of the film be sufficiently dense for thinning to occur slowly. But, on the other hand, it is also necessary that the surface tension be relatively weak so that it cannot overcome the opposite resistance of viscosity. Liquids having a high superficial viscosity and a relatively strong surface tension at the same time do not form bubbles. Liquids with a very low superficial viscosity do not form bubbles because the films surrounding them may become extremely thin in a short time, and an external or internal factor may break them. The chances for rupture are considerably less for a liquid having a low viscosity, as the bubbles are larger.

Plateau's principles further indicate that the superficial layer of liquids has a viscosity independent of the interior mass. In certain liquids, the

superficial viscosity is greater than the internal viscosity, and sometimes to a considerable extent as, for example, in water and in a solution of saponin. In other liquids, on the contrary, the superficial viscosity is weaker than the internal viscosity, as in the case of turpentine and alcohol. The formation of a lamina is the result of cohesion and viscosity of a liquid. When a lamina develops it progresses in this manner because it is pushed upwards continuously by the air; hence it continues to rise. On the other hand, the liquid to which its edges are attracted cannot follow the lamina because of its weight. Therefore this liquid remains behind, but cohesion and viscosity interfere so that no rupture occurs between the lamina formed and the surrounding liquid, the lamina growing until the action exerted upon the outer part of the air bubble has produced its effect. Changes in physical conditions, such as an increase in the temperature of the liquid decreasing capillary pressure, cause the lamina to collapse under a smaller weight. The foam duration of a 0.05 molar aqueous solution of isoamyl alcohol at 50°C. was found by Bartsch to be 32 per cent lower than that at 18°C.

When the superficial layers have a very high viscosity, molecular motion is difficult, with no changes occurring in the film before it is affected by surface tension. Surface tension, when sufficiently energetic, disrupts the film. If a bubble is formed from a glycerol solution and expanded, it gradually changes in thickness before it ruptures. These gradual changes in thickness are visible macroscopically, as indicated by the coloration of the film. In a glycerol solution, the course in color change is reversed, that is, from red to green of the last order before rupture, owing to the fact that glycerol adsorbs moisture from the air. The action of weight tends to make the film thinner, adsorption on the contrary having a tendency to make it thicker. The course of color change shows that thinning has at first the greatest effect, but that later thickening dominates. There is an interval in the existence of a bubble where these two factors are counter-balanced, that is, a period in which the film gains as much as it loses. In case a glycerol solution has a density not exceeding 1.1, it undergoes only a small decrease in density compensated by an increase in fluidity as a result of dilution with water. Changes in density are not the principal reason for this thinning.

When the film has retained a large proportion of water, the relation between superficial viscosity and surface tension must be sufficiently great to make rupture difficult. On the other hand, because of continued adsorption of water vapor, the film should not become thicker in any phase of its life. These two factors constitute the resistance of the bubble to rupture. Furthermore, if the ratio between the superficial viscosity and the surface tension increases, but not in proportion with viscosity, it



becomes insufficient to maintain permanency and the bubbles of the foam system collapse. This relation was pointed out by Plateau's (44) experiments (table 4).

The thinning of a film may be due to small concave surfaces formed along its contour. Thinning is slower for the same contour if the film occupies a larger surface. In the case of equal surfaces, thinning is less. This thinning rate increases with the diameter, but this is not the only factor involved in rupture, because if it were, the resistance of bubbles would increase with size. Usually these two influences compensate each other. A film of a given dimension has the greatest resistance to rupture when it is attached by its contour horizontally in a plane free from evaporation, agitation of air, or vibrations produced by the liquid. Two opposite

TABLE 4  
*Plateau's experiments*

SUBSTANCE	SUPERFICIAL VISCOSITY	FILM TENSION	VISCOSITY-TENSION RELATION
Water.....	100.00	14.60	6.85
Glycerol.....	60.42	8.00	7.55
Saturated solution of sodium carbonate...	91.14	8.56	10.65
Saturated solution of potassium nitrate...	96.35	11.22	8.59
Saturated solution of calcium chloride....	90.62	11.06	8.19
Marceille soap solution, 1/40.....	94.79	5.64	16.81
Domestic soap solution, 1/30.....	96.35	6.44	14.96
Potassium colophane solution (resin soap obtained by distillation of turpentine)..	84.89	7.68	11.05
Saponin solution, 1/100.....	Extremely high	8.74	Very high (not determined)
Albumin solution.....	Identical	11.42	Identical

forces act in a bubble, one centripetal, situated in the colorless part of the film, the other centrifugal, residing in the colored portion of the film with a maximum in the center corresponding to the so-called "black spot."<sup>2</sup> But apart from any action of the atmosphere, the interior of a fluid film is generally subjected to a continuous diminution in thickness by the internal current due to gravity and suction at the edges; sooner or later the interior ceases to possess the properties of matter as a whole. The film then becomes unstable with respect to the thinnest parts, which tend to become still thinner, owing to an apparent attraction between the surfaces of the film, discernible at its extremities, but becoming obvious when the thickness of the film is sufficiently reduced.

<sup>2</sup> The black spots never have a thickness (20–40m $\mu$ ) corresponding to the gray of the first order. The smallest thickness ever determined for these black spots was 100m $\mu$  (white of the first order); ordinarily, colors of higher orders are observed (300m $\mu$ ).

Rupture of a film for many liquids cannot always be determined because it is rather sudden, and the intermediate phases of rupture cannot always be observed. However, there are methods for determining the thinning of laminae and for measuring the limiting thickness, which is a measure of the stability of films composing a foam: (1) by tracing the causes of thinning—gravity and surface tension, and (2) by observing the decrease in thickness produced by a centrifugal force, as shown by Rickenbacher (55).

Upon a drum rotated around a vertical axis, Rickenbacher spread a film in a horizontal plane and observed the appearance of Newtonian colors, which served as a measure for the decrease in thickness. (The Newtonian interference colors depend upon the interference of light reflected from the anterior and posterior surfaces of the laminae.) By a comparison of the colors observed with Boys' (6) table of colors, the apparent thickness of the lamina with an angle of incidence may be determined (the more light reflected, the greater the angle of incidence). Boys' table of factors may also be used in converting apparent thickness into actual thickness.

The process of thinning may be observed to a limited extent in soap-water films. It is apparent that a film whose coloration approaches black exhibits marked instability. The continuous change in color is interrupted only when the black spots make their appearance, spreading rapidly. Decomposition of soap solutions, rendering them opaque, is another reason for film rupture. Precipitation, not an expansion of the film, occurs during the formation of these black spots, and may be attributed to the film's greater thickness at their edge. Black spots apparently are many times thicker than the portion of the film from which they are formed. If the formation of black spots were due to an expansion of the film, it is evident that they would not be present. The thickening of the edge of the film cannot be accounted for by contraction. Expansion of the upper part of the film and contraction of the lower and thicker part would be resisted less by viscosity and favored more by gravity than such expansions and contractions producing the phenomenon described. A rapid formation of a thin spot by an internal current would induce an accumulation at the edge of the material forming the interior of the film.

Van der Willigen (70) states that films of alkali solutions are rather unstable because the black zone is formed from the alkali portion of the soap. However, Plateau believes it more probable that soap itself tends to be removed from the solution and out of the colored part of the film. In a much more concentrated solution, it forms the black zone. This opinion is based on Newton's experiments, which proved that dark blue always precedes black and is observed only when the solution contains a high proportion of soap. This led Plateau to assume that the difference in composition between the black zone and other parts of the film is small

and the decrease in thickness gradual. The exploding force of a bubble is more intense the greater the proportion of water because the bubble bursts sooner, disintegrating its components, while, at the same time, the force of contraction has been weakened to the same extent. Conversely, the more soap present the less the intensity and the more energetic the force of contraction. This is to be doubted, because the force of contraction is surface tension, which decreases on increase in the proportion of soap. Leidenfrost (32) assumed as a consequence, therefore, that the force of rupture originates in water and the force of contraction in soap, so that the resistance of the bubbles varies with the concentration of the soap. Gibbs (16) pointed out that the retardation of the process by which films grow thinner cannot be accounted for in the formation of black spots, but may be attributed to a passive resistance to the motion due to the very viscous or gelatinous condition of the film. Also, it is unlikely that the formation of a stable film would occur without the support of such a resistance when rendered unstable during the same process by the precipitation of its mass. On the other hand, gelatinous properties are marked in soap solutions containing more soap than the amount required for the formation of films. This is true to a remarkable degree in saponin solutions. In an ordinary soap solution, however, no superficial viscosity greater than that of pure water was found (44). The resistance to an internal current is not measured by the resistance to motion, and it is doubtful whether a retardation of the tendency toward an internal current between elements adjacent to a black spot, changing slightly in thickness with time, may prevent rupture of the film (16). In a thick film the increase in surface tension with expansion necessary for the stability of the film is due to the presence of an excess of soap, or its components, at the surface as compared with the interior of the film.

### C. CONCLUSIONS

1. A film undergoes continuous decrease in thickness before its collapse. The thinning effect is produced by the weight of the lamina, as well as by the difference in the pressure at the terminal edge of the film and its interior.
2. Evaporation and condensation processes influence the transition to the unstable state, depending on the volatility of the components decreasing surface tension. Rupture is facilitated by volatility.
3. Decrease in the thickness of a film is accelerated or retarded, depending on whether the film is in a dry or a moist atmosphere.
4. An increase in temperature, which diminishes cohesion and capillary pressure, assists in the collapse of foam under a smaller weight.
5. The superficial viscosity of a film must be sufficiently high to make thinning of the film slow.

6. The surface tension should be relatively weak, so as not to overcome the opposing resistance of viscosity.

7. The ratio between superficial viscosity and surface tension is Plateau's criterion for the unstable state. When the ratio is larger and not proportional to viscosity, foam bubbles collapse.

8. Two forces act in a bubble—centripetal and centrifugal—with the centrifugal force located in the colored portion of the film and concentrated in the "black spot."

9. The intensity of rupture of a bubble depends on the concentration of soap.

## V. FOAM PREVENTION

### A. FACTORS INFLUENCING FOAM PREVENTION

In many industries foaming of liquids may or may not be desirable. As a matter of fact, in the heat treatment of certain liquids such as emulsified oils, foaming is a highly undesirable effect, and a number of methods are in use to prevent its taking place. In systems which foam, the general method for breaking or preventing the formation of foam is to have a superheating system above the evaporating surface of the liquid.

Pressure and high temperature reduce the tendency of liquid systems to foam. The lack of vapor space prevents substantial foam formation, as found in dehydrating emulsified crude oils (5). The properties of emulsified crude oils with respect to foaming are not additive. Intensity and persistency in foaming of emulsified oils to be dehydrated is decided by the nature of the oil. Foaming and the evaporation of water from the emulsified oil are markedly influenced by certain substances, such as gasoline. The foaming of emulsified oil increases as a function of the depth of the oil in the container, reaching its maximum at a definite temperature. Superheating and foaming occurring in the dehydration of emulsified oils produce evidence for the existence of a self-cooling effect.

Foaming of liquids often occurs during vacuum distillation. Jetting heated oil aids particularly in the dehydration of foaming oils (45). Small jets of emulsified oil are forced against a heated body beneath the surface of the liquid. Foam may be broken by passing it through a series of heated double grating plates (28). When passing the liquid to be distilled over a series of vertical plates, a countercurrent of the vapor prevents foaming (60). Foaming during steam distillation may be prevented by the addition of substances acting as defoamers. In most cases substances used in practice prevent or diminish foaming in a particular operation and are not universally applicable. The addition of many organic compounds to a foam made with a glue solution showed that in general fatty acids had a

TABLE 5  
*Examples of desirable foaming*

USE	FOAM AGENTS	REMARKS
Flotation process	Amyl alcohol } Cresol } foamers Xanthogenate } Oleic acid } "collectors" ad- Paraffin oil } sorbed by hy- Naphthylamine } drophobic ore Xylidine } powder	<p>Ore dispersed in water is replaced by substances suitable for foam formation; foam originates by blowing in air</p> <p>Hydrophobic ore particles not readily wetted adhere to the foam laminae</p> <p>By skimming off foam, the ore is separated from the hydrophilic admixture and the thoroughly wetted quartz and feldspar not incorporated in the foam, together with the aqueous phase, sink to the bottom</p>
Cleansing action of soap	Dirt particles	The dirt particles are carried by a hydrophobic skin in the interface as a result of emulsification, and are incorporated in the hydrophilic phase of the foam
Manufacture of beer, lemonade, etc.	Albumin Hop resins Saponin	The stability of a foam is a function of the fermentation process. It is assumed that acids originating in the fermentation process precipitate albumins and resins
Fire extinguishers	Solid particles (stabilizer)	A burning oil surface is extinguished by covering it with carbon dioxide and air containing foam; the foam contains in its laminae solid particles (stabilizer) separating the burning oil from the action of air. This action is similar to that of sand

greater effect upon the destruction of foam than the corresponding alcohols (38). No relation was found between the surface tension of the liquid and its ability to destroy foam, but in an homologous series the preventive effect appeared to increase with the molecular weight of the compound. With an increase in the concentration of the glue, the foaming power passed through a maximum. Certain mixtures were recommended for the prevention of foam; a mixture of linseed oil and alkali proved very effective. Petroleum, fat, milk, peanut oil, whale oil, oil of cloves, and various vege-

TABLE 6  
*Examples of undesirable foaming*

PROCESSES	FOAMING AGENTS	FOAM PREVENTERS	REMARKS
Paper production	Proteins Gelatin Dextrin Casein	Petroleum Fat Milk Peanut oil Oil of cloves Vegetable oils Alcohols of the fatty acid series	
Heat dehydration of crude oils	Introduction of steam into an emulsified oil undergoing dehydration, causing foaming	A mixture of linseed oil and alkali acts effectively in the prevention of foams	In the homologous series, the preventive effect increases with the molecular weight of the compound used
Heat dehydration of tars			Foaming in tars may occur when they contain less than 1 per cent of water

table oils were found to decrease foaming in practice. Removal of sludge has been very often proposed to inhibit foaming. The addition of lime to maintain the pH above the neutral point is also a means of preventing foam formation. Small amounts of chlorine added to sewage reduce its tendency to foam.

#### B. EXAMPLES OF DESIRABLE AND UNDESIRABLE FOAMING

Typical examples of the occurrence of desirable and undesirable foaming systems are illustrated in tables 5 and 6, respectively.

## VI. SUMMARY

Foams are labile formations. From a thermodynamic standpoint, they are produced with a great surface increase, and this increase acts against the tendency of the system to reach a stable state by diminution of the surface. Small surface tension combined with small vapor pressure characterizes a foam agent. Only capillary-active organic substances promote foam formation, because only liquids containing such substances fulfill these conditions. Concentration of the capillary-active substance used as a foam agent is of great importance. A decrease in tension and foam formation acts parallel only at certain concentration values specific for each capillary substance used. Foam formation ability decreases parallel with a decrease in surface activity of the component dissolved in water. The decrease in solubility limits the increase in surface activity. The intermediate members of the homologous series show the most suitable relation between solubility and surface activity (considerable solubility and at the same time a high surface activity). The influence of solubility itself is well expressed. An increase in solubility increases the foam formation ability. Surface tension changes cause adsorption, resulting in a Gibbs interface layer. The interface layer in foams is made up of various kinds of molecules. There is no foam without a heterogeneous interface layer, and homogeneous liquids are unable to produce stable foams. The purely molecularly dispersed members of an homologous series allow a greater heterogeneity than the partially colloiddally dissolved members. Solutions of surface-active substances foam better the more abrupt the concentration drop between the Gibbs layer and the solution.

Spreading is also a factor to be considered in foam formation. The rate of spreading is not parallel with the rate of surface increase in the formation of homogeneous layers, and the foam walls consist of heterogeneous layers.

The stability of foam systems is usually not high and is expressed in seconds of duration, although sometimes it reaches several days. Permanency may be increased under certain conditions. The coalescence of bubbles in a foam may be prevented by the addition of substances increasing the viscosity or rendering a certain elasticity which softens the skins of the interface. Foams may be classified as lyophilic and lyophobic with respect to their relation to electrolytes. Lyophilic foams have marked properties, such as viscosity and elasticity, and the influence of electrolytes on them is very small. Lyophobic foams are more stable and therefore less sensitive to electrolytes. Diminution of interfacial surface tension is one of the factors contributing to the stability of a foam system. Besides protective colloids, solids prevent coalescence of foam bubbles, maintaining a permanent foam system. The stabilization action of a solid in foam

systems depends upon the degree of dispersion of the solid and its nature, affinity, and concentration. The ability of a solid to wet is of great importance. Those substances less wetted by water adsorb the capillary-active substance from the aqueous solution forming the Gibbs interface layer. A relation exists between the ability to wet, adsorption, and flotation.

Changes occurring in the foam system and leading to rupture of the interface film or to its gradual thinning induce an unstable condition. Continuous diminution of the thickness of the interface film by internal currents is due to gravity and strong suction at the edges of the film. The effects of the motion of external molecules, evaporation, and condensation are important. Volatility influences the rupture of the superficial molecules of the film which are readily lost because of diminution in cohesion due to evaporation. A viscosity strong enough to slow down the thinning of the film prevents coalescence of bubbles. An increase in the temperature of a foaming liquid diminishes cohesion and also diminishes capillary pressure so that bubbles collapse under a smaller weight. The relation between superficial viscosity and surface tension should be large enough to make the rupture of the film difficult. The film is subjected to two different actions: (1) weight, which has a tendency to make it thinner, and (2) adsorption which, on the contrary, tends to make it thicker. Where two forces counterbalance, an equilibrium of the foam is reached. Foam systems present a typical case of a dynamic and not a static equilibrium. In the case of a solid used as a stabilizer, the dynamic equilibrium established between the rate of formation of the adsorbed layers of the foam-forming agent, or stabilizer, at the solid-liquid boundary on the one hand, and the rate of such layer formation at the gas-liquid interface on the other hand, determines to a certain extent the stability of a three-phase foam system. Small solid particles may contribute to the stabilization of a froth or foam, as in the case of mineralized froth, or in the ore flotation process, by preferential aggregation in the interface between two phases. Gas bubbles adhere to the solid particles in spite of the electrostatic repulsion of the components. The attraction between components of a foam system should be ascribed to more powerful capillary forces rather than to electrostatic forces. The function of capillary force is to separate the intermediate parts between two bubbles and to prevent coalescence where diminution of the surface occurs.

The adsorption layer in the interface evidently plays a most significant part in the stabilization of foam systems. In calculating the thickness of this layer, Bartsch proved that the maximum stability of these systems corresponds to a unimolecular film. An attempt has been made to explain the relation between the thickness of the interface layer and the stability of a foam system based on the Langmuir-Harkins theory of the fine struc-



ture of the boundary layers for those concentrations of foam-forming agents or stabilizers which lead to permanent foam systems. The intensity of adhesion between solid particles and gas bubbles is caused by the saturation of groups of molecules of the stabilizer rich in residual valences. These residual valences originate in the boundaries of the solid phase and are the result of the influence of the most saturated group of molecules of these substances, decreasing the surface tension at the gas-liquid boundaries, and the free boundary energy acquiring the smallest possible value. For small concentrations, the intensity of adhesion between solid particles and gas bubbles decreases with a decrease in the number of molecules of the foam agent or stabilizer causing adhesion.

At concentrations exceeding the condition of permanency, the adhesion between solid particles and gas bubbles occurs only when the dynamic equilibrium permits the formation of gas-liquid boundaries partially free from adsorbed molecules of the foam-forming agent or stabilizer.

#### REFERENCES

- (1) ADAMS AND MARCELIN: *Proc. Roy. Soc. London* **99**, 101, 103 (1921).
- (2) BARTSCH: *Kolloidchem. Beihefte* **20**, 1-51 (1925).
- (3) BARTSCH: *Kolloid-Z.* **38**, 178 (1926).
- (4) BECHOLD, DEDE, AND REINDERS: *Kolloid-Z.* **28**, 7 (1922).
- (5) BERKMAN, S., AND EGLOFF, G.: Superheating and Foaming Phenomenon in Dehydrating Emulsified Oils (unpublished).
- (6) BOYS: *Romance of Science Series*, Soc. for Promoting Christian Knowledge, London (1896); *Annual Report of the Smithsonian Institution*, p. 211, Washington, D. C. (1912); *Seifenblasen*, pp. 82 and 83, Leipzig (1913).
- (7) CARVER: *J. Am. Chem. Soc.* **45**, 63 (1923).
- (8) CHRISTIANSEN: *Ann. Physik* **40**, 107, 233 (1913).
- (9) DE COUDRES: Refer to L. Rhumbler, *Arch. Entw. Mech.* **7**, 225 (1898) and *Methodik der Entwicklungsmechanik*, p. 232, Urban and Schwarzenberg, Berlin (1923).
- (10) DONNAN: *Proc. Roy. Soc. London* **35A**, 557 (1911).
- (11) DONNAN: *Chem. Age* **31**, 282 (1923). *The Physical Chemistry of Interfaces*.
- (12) DUPRÉ: *Ann. chim. phys.* [4] **6**, **7**, **9**, **11**, and **14** (1865-1868). (5, 6 and 7 *Mémoire sur la théorie mécanique de la chaleur*.)
- (13) EVANS AND GEORGE: *Proc. Roy. Soc. London* **103A**, 190 (1923).
- (14) FELDMAN, A.: *Dissertation*, University of Chicago, 1921.  
HARKINS AND FELDMAN: *J. Am. Chem. Soc.* **44**, 2665 (1922).
- (15) FREUNDLICH, H.: *Kapillarchemie*, 3rd edition, p. 1091, chapter on foams. *Akademische Verlagsgesellschaft m.b.H.*, Leipzig (1923). Refer also to 4th edition (1930).
- (16) GIBBS, J. W.: *Study in Thermodynamics*. W. Engelmann, Leipzig (1892).
- (17) HAGEN: *Mém. de l'Acad. de Berlin* (1846); *Pogg. Ann.* **77**, 449 (1849).
- (18) HALL: *Proc. Roy. Soc. London* **9**, 56 (1895).
- (19) HARKINS, W. D.: *J. Am. Chem. Soc.* **42**, 700 (1920); *Colloid Symposium Monograph* **6**, 17-41 (1928).
- (21) HARKINS, CLARK, AND DAVIES: *J. Am. Chem. Soc.* **42**, 90 (1920); **39**, 354 (1917).

- (22) HELMHOLTZ AND PIOTROWSKY: Sitzber. Akad. Wiss. Wien, Math. naturw. Klasse 40, 607-58 (1860).
- (23) HENRY: Phil. Mag. 26, 541 (1845).
- (24) HISS: Dissertation, Heidelberg, 1913.
- (25) HOFFMAN: Z. physik. Chem. 83, 385 (1913).
- (26) HOUGH: J. Silliman [1] 17, 86 (1839).
- (27) LIIN: Physik. Z. 26, 497 (1925).
- (28) KNICHALIK, P.: German patent 522,782, December 4, 1929.
- (29) LAMARLE: Mém. acad. Belg. 5, 35-6 (1864). Sur la stabilité des systems liquides en lame minces.
- (30) LANGMUIR: J. Am. Chem. Soc. 40, 1361 (1918).
- (31) LANGMUIR: Proc. Nat. Acad. Sci. 3, 251 (1917); J. Am. Chem. Soc. 39, 1848 (1917); Trans. Faraday Soc. 15, 62 (1920).
- (32) LEIDENFROST: Tractatus, Duisburg, 1756. De aquae communis nonnullis qualitatibus tractatus.
- (33) LENARD: Ann. Physik 77, 463 (1915).
- (34) LEWIS, G. N.: Phil. Mag. [6] 15, 518, 705 (1908).
- (35) MARANGONI: Sull. espansione delle gocce d'un liquido galleggianti sulla superficie di altro liquido, Pavie (1865).
- (36) MARANGONI: Nuovo cimento [3] 3, 50, 97, 192 (1878); [2] 4 (1870); [2] 5-6 (1872).
- (37) MENSERUGGHE: Acad. Belg. 34 des mém. couronnés et Mém. des savants étrangers. Bull. acad. Belg. [2] 22, 308 (1866).
- (38) MEYER: Farben-Ztg. 481, 573, 619 (1930).
- (39) MICHAELIS AND RONA: Biochem. Z. 15, 196 (1909).
- (40) MILLNER: Phil. Mag. 13, 96 (1907).
- (41) MOSSOTTI: Lezioni elementari di Fisica Matematica, Florence, 1843.
- (42) MÜLFARTH: Ann. Physik 3, 328 (1900).
- (43) OSTWALD, WO., AND STEINER, A.: Kolloid-Z. 36, 342 (1925).
- (44) PLATEAU, JOSEPH: Statique expérimentale et théorique des liquides soumis aux seules forces moléculaires, Vol. I, chapter 5, pp. 93-4, 312-3, and chapter 6, p. 386. T. Gauthier Villars, Paris (1873).
- (45) PRUTZMAN, P. W.: U. S. patent 1,238,331, August 28, 1917.
- (46) QUINCKE, M.: Compt. rend. acad. Berlin, p. 132 (1868).
- (47) QUINCKE, M.: Wied. Ann. 35, 593 (1888).
- (48) RAMSDEN: Arch. Anat. Physiol., p. 517 (1894).
- (49) RAMSDEN: Proc. Roy. Soc. London 22, 156 (1903).
- (50) RAMSDEN: Trans. Faraday Soc. 22, 484 (1926).
- (51) RAYLEIGH, LORD: Lecture on The Tension of Recently Formed Liquid Surfaces, 1890.
- (52) RAYLEIGH, LORD: Scientific Papers 3, 92, 351, 397, 513, 572 (1887); Proc. Roy. Soc. London 48, 127 (1890); Chem. News 62, 17 (1890); Phil. Mag. [5] 33, 363 (1892); 48, 331 (1899).
- (53) REINDERS: Kolloid-Z. 13, 235 (1913).
- (54) REINDERS AND HOFFMAN: Z. physik. Chem. 83, 385 (1913).
- (55) RICKENBACHER: Kolloidchem. Beihefte 8, 139 (1916).
- (56) SEGNER: Commentar de figuris superficierum fluidarum, Vol. I, Göttingen, 1751.
- (57) SHORTER: Phil. Mag. 17, 560 (1909).
- (58) SHORTER: Phil. Mag. [6] 27, 718 (1914).
- (59) SHORTER: J. Soc. Dyers Colourists 31, 64 (1915).

- (60) SELIKER: German patent 506,545, February 27, 1925.
- (61) SZYSZKOWSKI: Z. physik. Chem. **64**, 385 (1908).
- (62) TAGGART: Phil. Mag. [6] **27**, 297 (1914); **28**, 364 (1914).
- (63) TALMUD AND SUCHOWOLSKAJA: Z. physik. Chem. **159**, 27-30 (1931).
- (64) THOMSON, J.: Application of Dynamics to Physics and Chemistry, pp. 226-7. Macmillan and Co., London (1888 and 1890).
- (65) TICKELL: Oil Gas J. **27**, N22, 175 (1928).
- (66) TRAUBE AND KICKE: Refer to Bartsch. reference 2, p. 47.
- (67) TRAUBE AND KLEIN: Kolloid-Z. **29**, 236 (1921).
- (68) TRAUBE AND NISHIZAWA: Kolloid-Z. **27**, 718 (1914).
- (69) VALENTINER: Physik. Z. **15**, 425 (1914).
- (70) VAN DER WILLIGEN: Pogg. Ann. **102**, 629 (1857).
- (71) WHITNEY AND STRAW: J. Am. Chem. Soc. **29**, 325 (1907).
- (72) WILSON: Phys. Rev. **16**, 8 (1920).
- (73) ZAWIDSKI: Z. physik. Chem. **35**, 77 (1900); **42**, 613 (1903).

# $\beta,\beta'$ -DICHLOROETHYL SULFIDE (MUSTARD GAS)

KIRBY E. JACKSON

*Department of Chemistry, Vanderbilt University, Nashville, Tennessee*

*Received September 27, 1933*

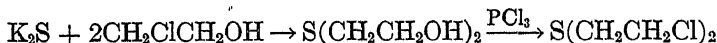
## I. HISTORY AND PREPARATION

The term "mustard gas," by which  $\beta,\beta'$ -dichloroethyl sulfide, or bis( $\beta$ -chloroethyl) sulfide, is most generally known, is an unfortunate one. By this name the most important chemical warfare agent is often confused with the true natural mustard oil which is prepared from black mustard seed and is chemically allyl isothiocyanate, having the following formula,  $\text{CH}_2:\text{CHCH}_2\cdot\text{N}:\text{C}:\text{S}$ . From a purely physical standpoint the classification "mustard gas" is also a misnomer, as the substance under ordinary conditions is a heavy oily liquid.

The early history of bis( $\beta$ -chloroethyl) sulfide is somewhat involved. The formation of a disagreeable smelling fluid by the action of ethylene upon chloride of sulfur had been observed by Despretz in 1822 (19), though its vesicant action was not noted. Thirty-seven years later Niemann (54) obtained a product by the action of ethylene upon "brown" chloride of sulfur (a mixture of sulfur monochloride and dichloride), which he found to be very toxic: "The most characteristic property of this oil is also a very dangerous one. It consists in the fact that the minutest traces which may accidentally come in contact with any portion of the skin, though at first causing no pain, produce in the course of a few hours a reddening and, on the following day, a severe blister, which suppurates for a long time and is very difficult to heal. Great care is therefore requisite in working with this compound." This product, for which the characteristics of "mustard gas" are so clearly set forth was, according to Niemann's analysis, the disulfide. It is thought that Guthrie in his examination of the action of chloride of sulfur upon ethylene (24) actually possessed some bis( $\beta$ -chloroethyl) sulfide, for he says: "Its smell is pungent and not unpleasant, resembling that of oil of mustard; its taste is astringent and similar to that of horseradish. The small quantities of vapor which it diffuses attack the thin parts of the skin, as between the fingers and around the eyes, destroying the epidermis. If allowed to remain in the liquid form on the skin, it raises a blister." Despite the fact that he ascribed to his product the composition of a disulfide and regarded it as different from the product of

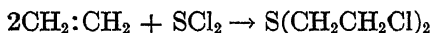
the reaction of ethylene with sulfur dichloride, his description of the compound as a pale yellow liquid of sp. gr. 1.346 at 19°C. and as highly toxic leaves no doubt but that he actually had some of the compound in hand.

V. Meyer (50) prepared bis( $\beta$ -chloroethyl) sulfide by allowing ethylene chlorohydrin to react with aqueous potassium sulfide and treating the resulting thiodiglycol with phosphorus trichloride:

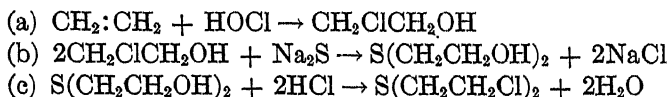


Clarke (13) improved the method by treating the intermediate, thiodiglycol, with concentrated hydrochloric acid on a water bath instead of with phosphorus trichloride, and Gomberg (22), in his studies, found that considerably less of the alkali sulfide could be used and that instead of 80 to 100 per cent of ethylene chlorohydrin a solution of 20 per cent gave as good results.

Gibson and Pope (21) prepared bis( $\beta$ -chloroethyl) sulfide by the direct addition of ethylene to sulfur dichloride, in accordance with the following equation:



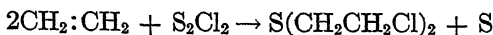
The ethylene, freed from alcohol vapor and thoroughly dried, is passed through the agitated sulfur dichloride, maintained at 40–45°C., at such a rate that some escapes unabsorbed. The ethylene is rapidly absorbed from the beginning, and the absorption ceases fairly suddenly when the reaction ends. This method was described in a report to the Chemical Warfare Service under date of January 16, 1918. It gives a better yield and is more expeditious than the method described by V. Meyer and improved by Clarke and Gomberg, which consists in converting (a) ethylene into ethylene chlorohydrin, (b) ethylene chlorohydrin into thiodiglycol by treatment with sodium sulfide, and (c) thiodiglycol into bis( $\beta$ -chloroethyl) sulfide by the action of hydrochloric acid, as shown in the following equations:



At the same time, the sulfur dichloride method is difficult to control, because bis( $\beta$ -chloroethyl) sulfide is acted on rapidly by sulfur dichloride; it is consequently necessary to work under such conditions that little or no sulfur dichloride remains long in contact with the bis( $\beta$ -chloroethyl) sulfide produced.

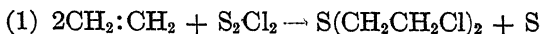
Later it was found that by using the monochloride and keeping the tem-

perature between 20°C. and 70°C. the yield of pure bis( $\beta$ -chloroethyl) sulfide varied between 90 and 98 per cent of that theoretically possible from the sulfur chloride:

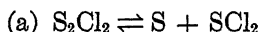


The day previous to the receipt of a cable from England to the effect that Gibson and Pope had successfully carried out the reaction, this experiment was started by Conant. The sulfur remaining after the distillation retained a mere trace of resinous matter, and no product other than bis( $\beta$ -chloroethyl) sulfide distilled. The reaction between ethylene and sulfur monochloride is accelerated by rise of temperature, absorption being about three times as rapid at 60°C. as at 30°C.

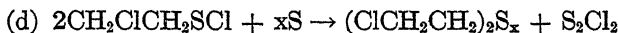
In place of the simple equation,



by which Gibson and Pope represented the reaction between ethylene and sulfur monochloride, Conant, Hartshorn, and Richardson (14) substituted the following:



and a side reaction,

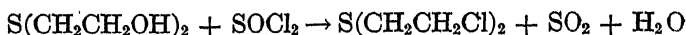


They found evidence favoring their view in determinations of the depression of the freezing point of bis( $\beta$ -chloroethyl) sulfide by partly ethylated sulfur monochloride, and concluded that reaction b is dominant until one molecule of ethylene has been absorbed for each one of sulfur monochloride originally present.

Others who prepared bis( $\beta$ -chloroethyl) sulfide by the same procedure obtained yields of only about 80 per cent. Mann, Pope, and Vernon (46) explained the discrepancy as follows: *when pure ethylene is absorbed by highly purified sulfur monochloride at 60°C., the yield is in the neighborhood of 80 per cent of the theoretical, but when the ethylene carries a small amount of alcohol vapor the yield rises to 99 per cent of the theory.* Insufficient scrubbing of the ethylene, prepared from alcohol, was thus responsible for the high yields.

Steinkopf, Herold, and Stahr (60) prepared bis( $\beta$ -chloroethyl) sulfide by allowing a chloroform solution of thiodiglycol to stand for one hour with thionyl chloride dissolved in chloroform; after fractionating, a colorless

product boiling at 108–109°C. at 15 mm. and melting at 10–12°C. was obtained.



Myers and Stephen (53) obtained bis( $\beta$ -chloroethyl) sulfide rapidly and continuously by spraying a mixture of 75 per cent sulfur dichloride and 25 per cent of the monochloride into an atmosphere of ethylene. By such a procedure further chlorination was avoided and the yield of bis( $\beta$ -chloroethyl) sulfide, melting at 8°C., was 93 per cent.

TABLE 1

*Variation in density and thermal expansion of bis( $\beta$ -chloroethyl) sulfide*

TEMPERATURE	DENSITY	VOLUME
<i>degrees C.</i>		<i>cc.</i>
15	1.2790	1.00000
20	1.2741	1.00324
25	1.2686	1.00824
30	1.2635	1.01226
35	1.2584	1.01642
40	1.2531	1.02067
45	1.2479	1.02497
50	1.2426	1.02901
55	1.2373	1.03372
60	1.2318	1.03830
65	1.2263	1.04295
70	1.2210	1.04747
75	1.2158	1.05193
80	1.2106	1.05648
85	1.2051	1.06133
88	1.2015	1.06452
90	1.1996	1.06614

## II. PHYSICAL PROPERTIES

### 1. Solubility in water; boiling point; melting point

Pure bis( $\beta$ -chloroethyl) sulfide is a colorless, neutral to litmus, almost odorless oil; however, Clarke (13) described it as smelling not unlike  $\alpha$ , $\epsilon$ -dichloropentane but far more unpleasant. It is almost insoluble in water, its solubility being 0.07 per cent at 10°C. (33), and 0.0043 moles per liter at 25°C. (68). Its boiling point has been reported as 97°C. at 10 mm. (30), 98°C. at 10 mm. (13), 104–105°C. at 13 mm. (40), 105°C. at 15 mm. (30), 107°C. at 15 mm. (51), 108–109°C. at 15 mm. (60), 120°C. at 34 mm., and 122.5°C. at 37 mm. At atmospheric pressure it decomposes at its boiling point, 217°C. (16). When poured into cold water it rapidly solidi-

fies in long colorless needles which melt over a range of several degrees, as determined by a number of observers: 8°C. (53), 10–12°C. (60), 13.5°C. (16, 67), 13.6°C. (61, 62), 13.82°C. (66), 14.4°C. (51).

### 2. Density and thermal expansion

Wilkinson and Wernlund (67) found by the use of a dilatometer tube the variation in density of bis( $\beta$ -chloroethyl) sulfide (m.p., 13.5°C.) and also the thermal expansion as shown in table 1. The variation in density per degree between 15°C. and 90°C. is 0.001058. The variation in volume of 1 gram per degree between the same temperature is 0.0006895 cc. The variation in volume in cubic centimeters per degree is 0.000881 cc.

TABLE 2

*Pressures produced by the action of sulfur monochloride on bis( $\beta$ -chloroethyl) sulfide*

TIME	PRESSURE PER 100 G. OF BIS( $\beta$ -CHLOROETHYL) SULFIDE		
	a	b	c
<i>hours</i>	<i>mm.</i>	<i>mm.</i>	<i>mm.</i>
0	0	0	0
16	64.5	273	328
19	71.8	312	368
22		336	409
25	95.4		
40	115.6	502	636
46	132.8	570	738
70	170.0		
72		791	944
89			944
96	170.0	753	924
136	160.0		
144		609	716
187	160.0	570	656

The curves for the variation in density and thermal expansion of bis( $\beta$ -chloroethyl) sulfide are both straight lines.

### 3. Pressure produced by the action of sulfur monochloride on bis( $\beta$ -chloroethyl) sulfide

Thompson and Kopp (62) determined the pressures produced by the action of relatively small concentrations of sulfur monochloride upon bis( $\beta$ -chloroethyl) sulfide. Column a of table 2 shows the data resulting from a mixture of 64 g. of the sulfide with 0.64 g. of sulfur monochloride (99:1 per cent); column b the data from a mixture of 41.1 g. of the sulfide with 1.37 g.



of the monochloride (96.78:3.22 per cent); column c shows the data from 49.4 g. of the sulfide with 2.47 g. of the monochloride (95.24:4.76 per cent).

The tables show that a maximum pressure is reached in approximately seventy-two hours, after which time a steady decrease is noted. When the apparatus was disconnected small amounts of hydrogen chloride and hydro-

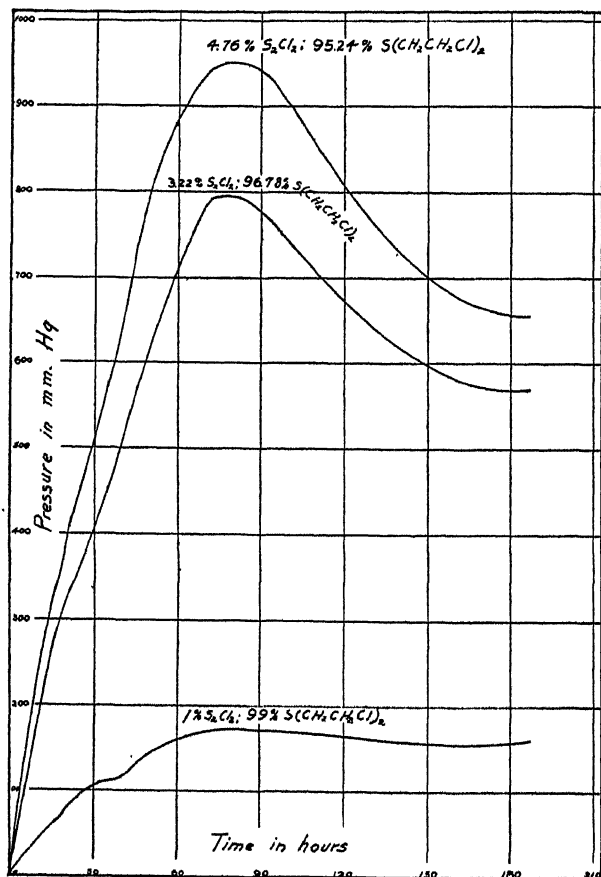


FIG. 1. PRESSURES PRODUCED BY ACTION OF SULFUR MONOCHLORIDE ON BIS( $\beta$ -CHLOROETHYL) SULFIDE

gen sulfide were detected in the gases liberated by the action of the monochloride on the sulfide. Owing to the solubility of hydrogen chloride in the latter substance, the increases in pressures shown in table 2 and the accompanying plot, figure 1, do not actually represent the total gases produced.

A secondary reaction takes place between the substances and the gases, as indicated by a decrease in pressure at the end of three days.

#### 4. The intersolubility of bis( $\beta$ -chloroethyl) sulfide and various solvents

The solubility of bis( $\beta$ -chloroethyl) sulfide in petroleum hydrocarbons was determined by Thompson and Odeen (63). The data obtained at various temperatures are given in tables 3, 4, 5, and 6 and illustrated by curves in figure 2.

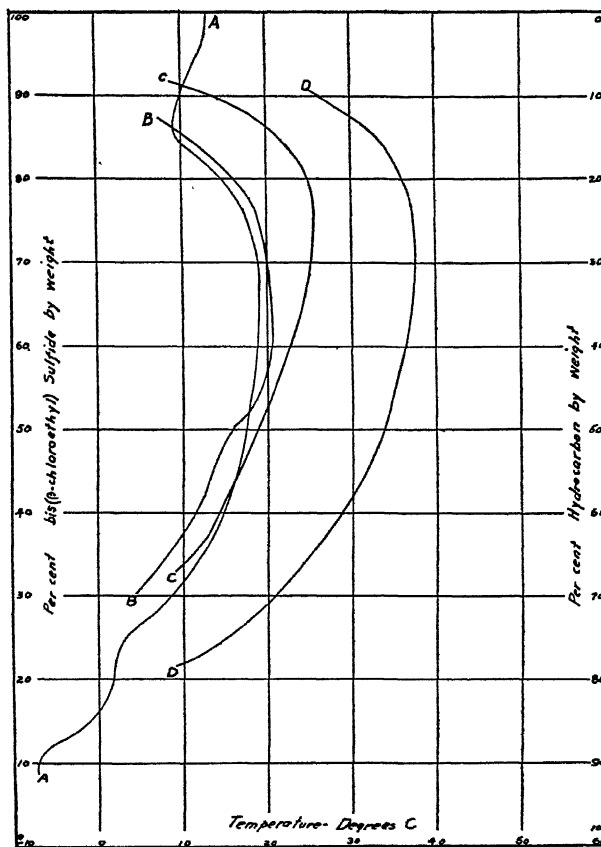


FIG. 2. INTERSOLUBILITY OF BIS( $\beta$ -CHLOROETHYL) SULFIDE AND HYDROCARBON SOLVENTS

Table 3 shows results obtained from a study of the intersolubility of bis( $\beta$ -chloroethyl) sulfide and ligroin (b.p. less than 100°C.) at various temperatures and illustrated by the curve AA in figure 2. The critical point of solubility, that is, the temperature above which the two substances would be mutually soluble in all proportions, was found to be 19°C.; the highest temperature at which solid bis( $\beta$ -chloroethyl) sulfide could exist in the presence of its liquid phase and a ligroin solution of it was 8.8°C.

The critical temperature of solubility of bis( $\beta$ -chloroethyl) in gasoline (b.p. 60–180°C.) was found to be 20.4°C. The data from their intersolubility is given in table 4 and illustrated by curve BB in figure 2.

Table 5 shows the results secured from a study of the mutual solubilities of bis( $\beta$ -chloroethyl) sulfide and kerosene (b.p. 160–260°C.) at various temperatures. These determinations indicate that the critical point of

TABLE 3  
*Intersolubility of bis( $\beta$ -chloroethyl) sulfide and ligroin*

TEMPERATURE	LIGROIN		BIS( $\beta$ -CHLOROETHYL) SULFIDE		SOLUBILITY	
					Ligroin	Bis( $\beta$ -chloroethyl) sulfide
°C.	cc.	grams	cc.	grams	per cent by weight	per cent by weight
13.0	0	0.00	20	25.44		100.00
13.0	1	0.67	20	25.44	2.6	97.4
12.0	2	1.34	20	25.44	4.9	95.1
10.0	4	2.67	20	25.44	9.5	90.5
9.5	5	3.34	20	25.44	11.6	88.4
9.0	6	4.00	20	25.44	13.6	86.4
9.6	7	4.67	20	25.44	15.5	84.5
12.0	8	5.34	20	25.44	17.4	82.6
14.0	9	6.00	20	25.44	19.1	80.9
17.22	12	8.01	20	25.44	23.9	76.1
19.0	19	12.69	20	25.44	33.3	66.7
19.0	24	16.03	20	25.44	38.6	61.4
18.3	29	19.36	20	25.44	43.2	56.8
17.5	39	26.04	20	25.44	50.6	49.4
16.0	49	32.72	20	25.44	56.3	43.7
14.5	59	39.39	20	25.44	60.7	39.3
10.3	79	52.75	20	25.44	67.5	32.5
6.5	99	66.10	20	25.44	72.2	27.8
2.5	119	79.46	20	25.44	75.8	24.2
0.2	30	20.03	3	3.82	83.9	16.1
−2.5	30	20.03	2.5	3.18	86.3	13.7
−6.5	30	20.03	2	2.54	88.7	11.3
−7.5	30	20.03	1.5	1.91	91.3	8.7

solubility of the sulfide and kerosene is 25.6°C. These data are illustrated in figure 2 by curve CC.

The results secured in experiments on the intersolubility of bis( $\beta$ -chloroethyl) sulfide and railroad light oil (b.p. 140–315°C.) are shown in table 6 and illustrated by the curve DD in figure 2. The critical point of solubility was found to be 37°C.

From these results it is evident that with the increase in the percentage

of high boiling hydrocarbons, the critical temperature of solubility for these solvents and bis( $\beta$ -chloroethyl) sulfide will also increase. This is very nicely illustrated by a comparison of the solubility curves as shown in figure 2.

TABLE 4  
*The intersolubility of bis( $\beta$ -chloroethyl) sulfide and gasoline*

TEMPERATURE	BIS( $\beta$ -CHLOROETHYL) SULFIDE		GASOLINE		SOLUBILITY	
					Bis( $\beta$ -chloroethyl) sulfide	Gasoline
°C.	cc.	grams	cc.	grams	per cent by weight	per cent by weight
7.3	20.05	25.43	5.00	3.686	87.34	12.66
18.0	20.05	25.43	10.00	7.371	77.53	22.47
20.4	20.05	25.43	17.75	13.095	66.02	33.98
18.7	20.05	25.43	30.00	22.113	53.49	46.51
14.0	20.05	25.43	40.00	29.484	46.32	53.68
9.5	20.05	25.43	60.00	44.226	36.51	63.49
4.5	20.05	25.43	80.00	58.968	30.14	69.86
7.8	15.00	19.03	3.80	2.801	87.17	12.83
13.5	15.00	19.03	30.05	22.150	46.21	53.79
9.0	15.00	19.03	45.00	33.170	36.46	63.54
4.2	15.00	19.03	60.00	44.226	30.08	69.92

TABLE 5  
*The mutual solubilities of bis( $\beta$ -chloroethyl) sulfide and kerosene*

TEMPERATURE	KEROSENE		BIS( $\beta$ -CHLOROETHYL) SULFIDE		SOLUBILITY	
					Kerosene	Bis( $\beta$ -chloroethyl) sulfide
°C.	cc.	grams	cc.	grams	per cent by weight	per cent by weight
14.2	60	48.24	25	31.88	60.21	39.79
9.5	80	64.32	25	31.88	66.87	33.13
21.9	7	5.63	25	31.88	15.07	84.93
24.3	9	7.23	25	31.88	18.58	81.42
25.6	12	9.65	25	31.88	23.33	76.67
25.6	13	10.45	25	31.88	24.79	75.21
25.6	14	11.26	25	31.88	26.20	73.80
25.0	18	14.47	25	31.72	31.33	68.67
14.3	6	4.82	33	41.88	10.32	89.68
8.9	6	4.82	43	54.47	8.12	91.88

Thompson, Black, and Sohl (61) determined the intersolubility of bis( $\beta$ -chloroethyl) sulfide (double-distilled at 13–14 mm. and melting at 13.6°C.) and ethyl alcohol. The solubilities of the alcohol and the sulfide

were determined by heating the two substances together in the proportions given in the table until a clear solution was produced. The temperatures at these points were noted. Upon cooling the solution very slowly a characteristic cloudiness appeared, resulting from the formation of two liquid phases. The temperatures at which the cloudiness first appeared were noted. The mutual solubilities of the sulfide and 92.5 per cent alcohol are given in table 7 and the solubilities of the sulfide and absolute alcohol are given in table 8, the data from each being illustrated by the solubility curves in figure 3. The temperature given in the tables, in each case, is the mean obtained from the temperature at which the solution cleared and that at which the cloudiness was produced. Above the

TABLE 6

*The intersolubility of bis( $\beta$ -chloroethyl) sulfide and railroad light oil*

TEMPERATURE	RAILROAD LIGHT OIL		BIS( $\beta$ -CHLOROETHYL) SULFIDE		SOLUBILITY	
					Railroad light oil	Bis( $\beta$ -chloro- ethyl) sulfide
$^{\circ}\text{C.}$	cc.	grams	cc.	grams	per cent by weight	per cent by weight
25.0	2.45	1.926	14.95	18.966	9.22	90.78
28.0	5.12	4.024	25.00	31.715	11.26	88.74
31.3	3.74	2.940	14.95	18.966	13.42	86.58
37.0	13.69	10.760	14.95	18.966	36.20	63.80
35.0	20.19	15.869	14.95	18.966	45.52	54.48
33.0	26.19	20.585	14.95	18.966	52.05	47.95
23.6	38.57	30.316	14.95	18.966	61.51	38.49
20.9	56.62	44.503	14.95	18.966	70.12	29.88
14.5	75.07	59.005	14.95	18.966	75.67	24.33
9.3	87.42	68.692	14.95	18.966	78.36	21.64

curves in figure 3 a homogeneous system exists and below the curves there is a heterogeneous system of the two liquid phases.

The effect of water upon the critical point of solubility is very decided. With 92.5 per cent alcohol a critical temperature of solubility of  $38.6^{\circ}\text{C.}$  was secured, but with absolute alcohol the critical temperature was  $15.6^{\circ}\text{C.}$

*5. Freezing point constant; vapor pressure; heat of combustion; refractive index*

Delépine, Fleury, and Ville (18) determined the freezing point constant,  $K$ , of bis( $\beta$ -chloroethyl) sulfide. Raoult's law gives  $K = 60 - 2$ , but this value diminishes very rapidly with moderate increases of the concentration. When Louguinine and DuPont's modified formula (42),  $dt = Kn/(n + N)$  is applied, better agreement is realized and  $K = 38$ .

Herbst reports the measurements, by an air-saturation method, of the vapor pressure of an impure specimen of bis( $\beta$ -chloroethyl) sulfide (m.p.  $12.85^{\circ}\text{C}.$ ) at  $22^{\circ}\text{C}.$  and  $39^{\circ}\text{C}.$  as 0.087 and 0.358, respectively. Apart from this, there appears to be no record of the determination of the vapor pressure of this substance at ordinary temperatures until a study was made by Mumford, Phillips, and Ball (52), though a number of values are quoted in the literature. These, however, are very discordant, ranging from 0.01 mm. (55) to 0.051 mm. (34) at  $10^{\circ}\text{C}.$ , and from 0.06 mm. (55) to 0.11 mm. (64) at  $20^{\circ}\text{C}.$

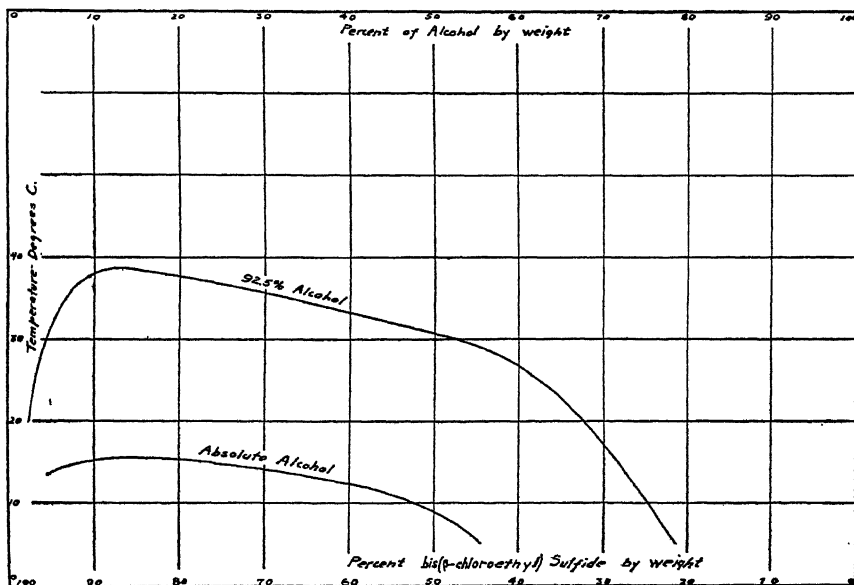


FIG. 3. INTERMISCIBILITY OF ALCOHOL AND BIS( $\beta$ -CHLOROETHYL) SULFIDE

The results of the vapor pressure determinations, represented by the equation,

$$\log_{10} p = 8.3937 - 2734.5/T$$

are given in table 9 and illustrated by graph in figure 4.

The heat of combustion of liquid bis( $\beta$ -chloroethyl) sulfide in the calorimetric bomb was determined as 743.3 Calories at constant volume by Heycock and Mills (31). They also determined its heat of formation from rhombic sulfur, amorphous carbon, and gaseous hydrogen and chlorine as 67.9 Calories.

Refractive indices were determined by Palmer (56) as 1.52776 for  $\text{H}_2\text{O}$ ,

1.53125 for  $\text{Na}_2\text{D}$ , and 1.53999 for  $\text{H}_\text{F}$  at  $15^\circ\text{C}$ . The molecular depression of the freezing point by benzene was determined as 54.6; the latent heat of fusion is hence calculated as about 30 Calories.

### 6. Surface tension determinations

Harkins and Ewing (26) made the surface tension determinations between bis( $\beta$ -chloroethyl) sulfide and various aqueous solutions by the drop-

TABLE 7

*Mutual solubility of bis( $\beta$ -chloroethyl) sulfide and 92.5 per cent alcohol*

TEMPERATURE	BIS( $\beta$ -CHLOROETHYL) SULFIDE	92.5 PER CENT ALCOHOL	SOLUBILITY	
			Bis( $\beta$ -chloroethyl) sulfide	92.5 per cent alcohol
$^\circ\text{C}$ .	grams	grams	per cent by weight	per cent by weight
19.9	31.97	0.81	97.53	2.47
30.9	31.97	1.62	95.18	4.82
35.7	31.97	2.43	92.93	7.07
37.8	31.97	3.24	90.80	9.20
38.4	31.97	4.05	88.75	11.25
38.5	31.97	4.86	86.80	13.20
38.6	31.97	5.67	84.93	15.07
37.9	31.97	7.29	81.42	18.58
37.6	31.97	8.91	78.20	21.80
36.7	31.97	10.53	75.22	24.76
35.8	31.97	12.15	72.45	27.55
35.2	31.97	14.58	68.67	31.33
34.4	31.97	17.02	65.27	34.73
33.6	31.97	19.45	62.18	37.82
33.0	31.97	21.88	59.37	40.63
32.5	31.97	24.31	56.76	43.24
31.0	31.97	26.74	54.45	45.55
31.4	31.97	29.17	52.29	47.71
30.6	31.97	32.42	49.66	50.34
29.9	28.13	32.42	46.47	53.53
28.3	24.30	32.42	42.84	57.16
25.6	20.46	32.42	38.70	61.30
21.7	16.62	32.42	33.91	66.09
17.6	14.06	32.42	30.26	69.74
12.8	11.51	32.42	26.20	73.80
5.1	8.95	32.42	21.64	78.36

weight method, using the proper precautions as determined by Harkins and Brown (25). The radius of the tip used was 0.24995 cm. for all the interfacial work, except when a 0.1 per cent solution of turkey-red oil was used in the absence of sodium hydroxide, when the radius was 0.275 cm. For bis( $\beta$ -chloroethyl) sulfide alone a tip of 0.27441 cm. radius was used.

The density of this latter liquid was 1.2732 g. per cubic centimeter. In all of the experiments the aqueous solution was dropped upward through this liquid without allowing time for mixing. The values for  $r/a$  and  $r/v^{\frac{1}{3}}$  indicate the shape of the drop (25).

TABLE 8  
*Mutual solubility of bis( $\beta$ -chloroethyl) sulfide and absolute alcohol*

TEMPERATURE	BIS( $\beta$ -CHLOROETHYL) SULFIDE	ABSOLUTE ALCOHOL	SOLUBILITY	
			Bis( $\beta$ -chloroethyl) sulfide	Absolute alcohol
$^{\circ}\text{C.}$	grams	grams	per cent by weight	per cent by weight
13.6	31.9	1.57	95.31	4.69
14.5	31.9	2.35	93.13	6.87
14.8	31.9	3.14	91.04	8.96
15.3	31.9	4.71	87.13	12.87
15.6	31.9	6.48	83.12	16.88
15.5	31.9	7.85	80.25	19.75
14.8	31.9	10.20	75.76	24.24
14.2	31.9	12.56	71.75	28.25
13.6	31.9	15.70	67.00	33.00
12.2	31.9	19.62	61.91	38.09
11.8	31.9	23.55	57.52	42.48
10.6	31.9	27.47	53.73	46.27
9.1	31.9	31.40	50.39	49.61
7.5	31.9	35.32	47.45	52.55
5.6	31.9	39.25	44.83	55.17

TABLE 9  
*Experimental vapor pressure of bis( $\beta$ -chloroethyl) sulfide*

TEMPERATURE	MEAN PRESSURE	MAXIMUM DEVIATION FROM MEAN, $\pm$	VAPOR PRESSURE (CALCULATED)
$^{\circ}\text{C.}$	mm.		
0	0.035	0.007	0.024
10	0.055	0.005	0.054
15	0.075	0.008	0.079
20	0.115	0.007	0.115
30	0.225	0.025	0.230
40	0.450	0.050	0.450
50	0.830	0.100	0.850
60	1.550	0.150	1.520

The surface tension at  $20^{\circ}\text{C.}$  between bis( $\beta$ -chloroethyl) sulfide and water is 28.36 dynes per centimeter, or about 5 dynes less than that between benzene and water. The surface tension of this liquid toward its vapor is



42.82 dynes per centimeter, which is 14 dynes higher than that of benzene. One of the most important characteristics of a liquid in this connection is the amount of work done when 1 sq. cm. of the surface of the liquid approaches 1 sq. cm. of a water surface. This value is 87.26 ergs in the case of bis( $\beta$ -chloroethyl) sulfide, which indicates that this liquid, while only

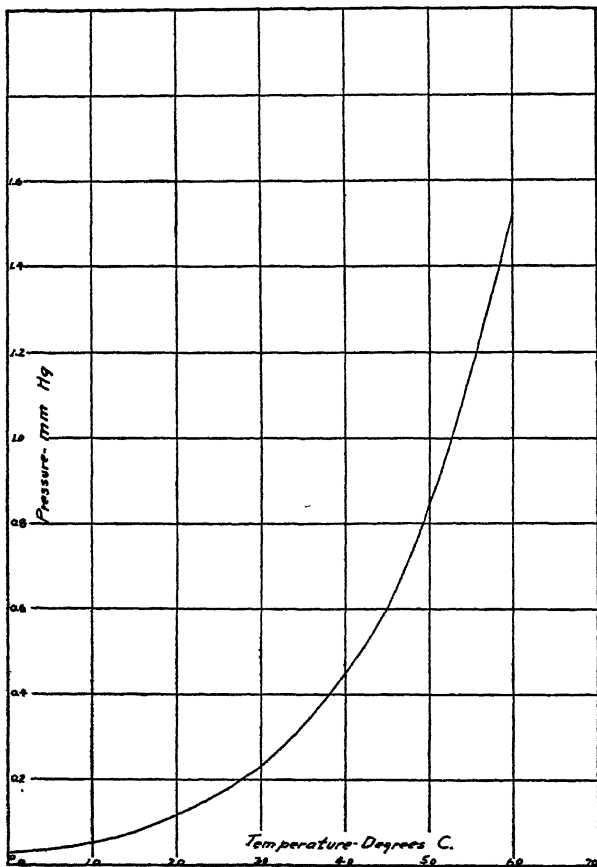


FIG. 4. VAPOR PRESSURE OF BIS( $\beta$ -CHLOROETHYL) SULFIDE

very slightly soluble in water, has a high attraction for it, although the magnitude of the work done is slightly less than that found for alcohol.

#### 7. The compressibility of bis( $\beta$ -chloroethyl) sulfide

The compressibility of bis( $\beta$ -chloroethyl) sulfide was determined by Adams and Williamson (2) after the method of Adams, Williamson, and Johnson (3). The sulfide was surrounded by kerosene in the interior of a

thick-walled steel cylinder, one end being closed while the other was fitted with a piston and leak-proof packing. By means of a hydraulic press the piston was forced into the cylinder, thus subjecting the kerosene and the sulfide to hydrostatic pressure and thereby decreasing the volume of each.

TABLE 10

*The surface tension between bis( $\beta$ -chloroethyl) sulfide and various aqueous solutions*

SUBSTANCE	VOLUME OF 1 DROP	DENSITY OF SUBSTANCE	$r/a$	SURFACE TENSION
	cc.			degrees per cm.
Vapor.....	0.03555		1.04	42.82
Water.....	0.10620	0.9982	0.595	28.36
0.1 <i>N</i> hydrochloric acid.....	0.10900	1.00011	0.5385	28.90
0.1 <i>N</i> sodium hydroxide.....	0.04612	1.0032	0.8066	12.78
0.1 <i>N</i> sodium carbonate.....	0.06950	1.0025	0.6635	18.82
1 per cent turkey-red oil.....	0.05670	1.0000	0.6070	14.47
1 per cent turkey-red oil in 0.1 <i>N</i> sodium carbonate.....	0.02977	1.0035	1.0000	8.35
			$r/r^{\frac{1}{2}}$	
1 per cent Twitchell's solution.....	0.04384	0.99924	0.7088	12.32
1 per cent Twitchell's solution in 0.1 <i>N</i> sodium carbonate.....	0.04663	1.00311	0.6944	12.89
1 per cent corn oil solution.....	0.45770	0.99840	0.6980	12.94
1 per cent corn oil solution in 0.1 <i>N</i> sodium carbonate.....	0.03936	1.00360	0.7350	10.91

TABLE 11

*Decrease in volume of bis( $\beta$ -chloroethyl) sulfide under pressure*

PRESSURE	$-\Delta v/v_0$ (OBSERVED)		$-\Delta v/v_0$ (CALCULATED)
	1	2	
<i>megabars</i>			
392	0.0000	0.0000	0.0000
840	0.0175	0.0179	0.0177
1280	0.0327	0.0331	0.0326
1713	0.0447	0.0450	0.0450

The decrease in the volume of the sulfide for each increment of pressure was determined from a pair of readings of pressure and corresponding piston displacement. Pressures were measured by an electric method, the precision being such that one scale division was equal to about one megabar; the displacement of the piston was read on a dial micrometer

which was graduated to 0.01 mm. Temperature regulation was obtained by an electric heating coil of asbestos-covered "nichrome" wire wrapped about the heavy steel cylinder and covered with an insulating layer of felt. About 10 g. of bis( $\beta$ -chloroethyl) sulfide was contained in a steel capsule closed at the top and terminating in a capillary which dipped into a cup containing mercury. Pressure could thus be transmitted through the mercury seal to the interior of the capsule.

Table 11 shows the results for the decrease in volume at 31.5°C. of the bis( $\beta$ -chloroethyl) sulfide; the second and third columns of the table give for two separate runs the value of  $-\Delta v/v_0$ , where  $\Delta v$  is reckoned from the initial pressure,  $P_0 = 392$  megabars, and  $v_0$  is the volume of the liquid at 31.5°C. and atmospheric pressure. The value of  $\Delta v$  as a function of  $P$  may be represented by a power series yielding the equation

$$\Delta v/v_0 = 4.24 \times 10^{-5} (P - P_0) - 6.3 \times 10^{-5} (P - P_0)^2$$

However, the results are expressed equally well by the exponential equation,

$$-\Delta v/v_0 = 0.118 [I - e^{-0.364 \times 10^{-3} (P - P_0)}] \quad (1)$$

which gives a more reasonable course to the compressibility curve and hence is to be preferred for extrapolating to zero pressure. In the third column are shown the values of  $-\Delta v/v_0$  calculated from equation 1.

By differentiation this equation becomes

$$-dv/dP = 49.5e^{-0.364 \times 10^{-3} P} \quad (2)$$

from which the compressibility ( $-dv/dP$ ) at  $P = 0$  was found to be  $495 \times 10^{-6}$  per megabar, while at 1000 and 2000 megabars, respectively, the compressibility is  $34.4 \times 10^{-6}$  and  $23.9 \times 10^{-6}$ .

### 8. Freezing pressure; freezing point-solubility curve

After the conclusion of the measurements of compressibility the freezing pressure and resultant change of volume at a few temperatures were determined; this could be done without removing the material from the apparatus. The desired quantities may be obtained from a series of readings, at constant temperature, of pressure  $P$  and piston-displacement  $R$ . When freezing or melting of the substance in the capsule takes place,  $P$  remains constant while  $R$  increases or decreases, and the resulting discontinuity at once locates the freezing pressure for the given temperature.

It was found that the necessary readings had to be taken with decreasing pressure since, on account of the propensity of liquids for undercooling to a temperature several degrees below their melting point, they will generally support a pressure far beyond the true freezing pressure before solidifica-

tion takes place. Bis( $\beta$ -chloroethyl) sulfide requires at ordinary temperatures about 1000 megabars superpressure to start it freezing, which corresponds to an undercooling of 14°C.

The results for the freezing points at several pressures are shown in table 12, which also gives the change of volume  $V_l - V_s$  in cubic centimeters per gram. From these results it may be seen that the compressibility measurements of table 11 were extended by about 500 megabars into the region of undercooled liquid.

In order to test for the possible existence of other solid modifications of this substance, the pressure on it was increased to 12,000 megabars at 38°C., but no new forms were discovered.

TABLE 12

*Results for freezing pressures, and change of volume on freezing*

TEMPERATURE	FREEZING PRESSURE	$V_l - V_s$	$dP/dT$	$\Delta H$ (CALCULATED)
°C.	<i>megabars</i>	<i>cc. per gram</i>	<i>megabars per degree</i>	
13.9	1	(0.054)*	68	
21.9	570	0.050	71	25
29.6	1110			
31.4	1210	0.047	74	25
38.9	1800	0.042	77	25

\* By extrapolation.

From the data contained in table 12 the latent heat of melting was calculated by the Clausius-Clapeyron equation. Thus

$$\Delta H = 0.02391T (V_l - V_s) dP/dT$$

where  $\Delta H$  is the latent heat in calories per gram,  $T$  is the absolute temperature,  $V_l - V_s$  is the change of volume on melting in cubic centimeters per gram, and  $dP/dT$  is the slope of the  $P - T$  curve expressed in megabars per degree; practically it is the pressure in megabars required to raise the melting point one degree. Over the range of temperature covered by the measurements of Adams and Williamson it was found that  $\Delta H = 25$  calories per gram. This is a value which is higher than the latent heat of melting of most substances but, on the other hand, the compressibility and change of volume upon freezing do not differ markedly from the average for organic liquids.

Wilkinson, Neilson, and Wylde (66) in determining the freezing point solubility curve for bis( $\beta$ -chloroethyl) sulfide used a product melting at 13.82°C., high-grade roll sulfur, and especially purified rhombic, monoclinic,

and amorphous sulfur. These determinations were made by mixing equal amounts of the two substances in a test tube and heating slowly (2°C. per 5 minutes), noting the temperature at which the two phases become one, then cooling and observing the temperature at which the second phase appeared. These results are shown graphically in figure 5.

TABLE 13  
*Solubility of roll sulfur*

SULFUR	NATURE OF THE SOLID PHASE AT GIVEN TEMPERATURE		
	Solid bis( $\beta$ -chloroethyl) sulfide	Solid sulfur	Liquid phase
<i>per cent</i>	°C.	°C.	°C.
0.00	13.82		
0.10	13.79		
0.30	13.75		
0.70	13.70		
0.90	13.66		
1.48		24.0	
2.44		43.0	
3.38		54.0	
4.31		61.0	
6.10		74.0	
9.09		85.0	
16.66		99.0	
20.00		102.5	
25.18			110.0
31.00			120.0
32.00			122.0
35.00			127.0
40.00			133.5
60.00			143.5
78.00			135.0
85.00			124.0
87.00			117.0
90.00			105.0
94.00		108.0	
100.00		114.0	

Rhombic and monoclinic sulfur were found to have the same solubility, while amorphous sulfur had a solubility of less than 0.14 per cent below 120°C.; below 78°C. the solid phase which separated was rhombic sulfur, while above this temperature it was monoclinic.

Attempts to determine accurately the partition coefficient of bis( $\beta$ -chloroethyl) sulfide have not been highly successful, owing to the rapidity with which it hydrolyzes, but using xylene and water at 20°C. Lynch, Smith, and Marshall (43) determined it as over 200.

## III. CHEMICAL PROPERTIES

## 1. Action with metals

Aluminum, lead, brass, iron, bronze, zinc, and tin are not acted on by bis( $\beta$ -chloroethyl) sulfide at ordinary temperatures; with the first three metals no appreciable action is noted at 100°C., but the others named are attacked at this temperature, the action being most rapid in the case of tin. Titanic and stannic chlorides yield additive compounds with the substance. Sulfur dichloride acts vigorously on bis( $\beta$ -chloroethyl) sulfide at ordinary temperatures (21).

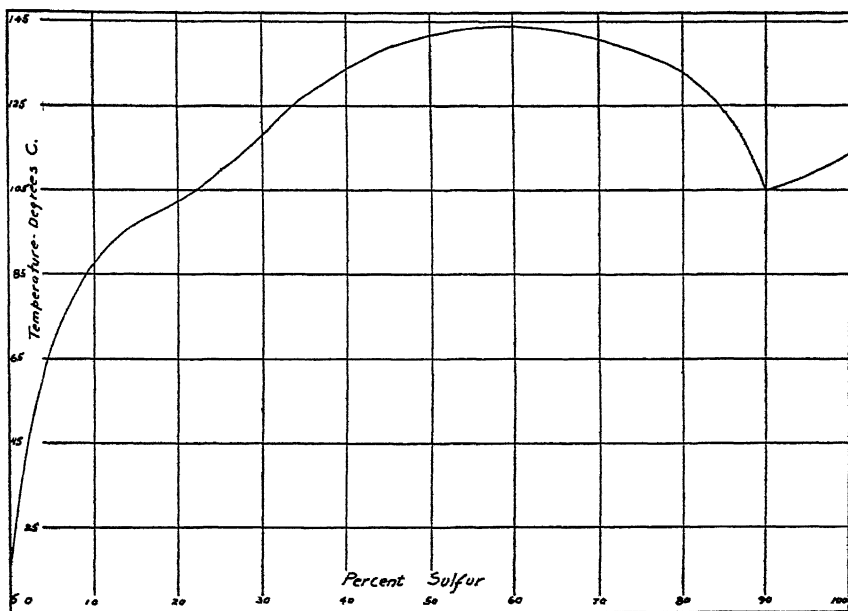
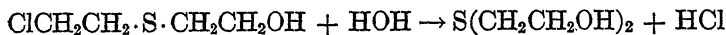


FIG. 5. SOLUBILITY-FREEZING POINT CURVE OF SULFUR IN BIS( $\beta$ -CHLOROETHYL) SULFIDE

Kretov (36) investigated the interaction between bis( $\beta$ -chloroethyl) sulfide and zinc dust and found it to be a complex reaction resulting in the formation of a large number of compounds. The sulfide is entirely decomposed with formation of ethylene, hydrogen sulfide, diethylene disulfide, zinc chloride, divinyl sulfide, diethyl sulfide, and ethyl vinyl sulfide. The decomposition is caused by metallic zinc and not by zinc oxide present as an impurity. The principal products of the reaction between bis( $\beta$ -chloroethyl) sulfide and zinc in the presence of alcohol are esters of thiodiglycol of the general formula  $S(CH_2CH_2OR)_2$ , which are stable liquid compounds, capable of distillation under atmospheric pressure without decomposition.

## 2. Hydrolysis

It was generally assumed and later proved (57) that the hydrolysis of bis( $\beta$ -chloroethyl) sulfide occurs in two stages:



Wilson, Fuller, and Schur (68) showed that the hydrolysis of bis( $\beta$ -chloroethyl) sulfide in contact with water is a two-stage reaction which affects only the molecules dissolved in the aqueous phase.

The rate of the first stage of the reaction determines that of the second. Once steady conditions are established, the concentration of the intermediate compound automatically increases or decreases, keeping the rate of the second stage just equal to that of the first. The amount of intermediate compound present at any time is extremely small, and the time required to build up the equilibrium concentration is generally negligible.

The simplest expression for the rate of hydrolysis at 25°C. in alkaline solutions, where there is no tendency toward reversal, is:

$$dC/dt = 2K_1 - (M) = 0.304 (M)$$

where  $dC/dt$  is the rate of hydrolysis in millimoles of hydrogen chloride produced per liter per minute and  $(M)$  is the concentration of bis( $\beta$ -chloroethyl) sulfide in millimoles per liter.

The second stage of the hydrolysis reaction is substantially non-reversible except in very strong hydrochloric acid solutions. The first stage is, however, reversible to a considerable extent in the presence of even dilute acids. This reversibility of one stage serves to retard but not to stop the reaction, because the intermediate hydrolysis compound is being continually destroyed by the irreversible second stage.

The mathematical expression for the rate of hydrolysis in acid solutions is, therefore,

$$dC/dt = K_1(M) - k_1(I)(\text{H}^+)(\text{Cl}^-) + K_2^1(I)(\text{H}^+)$$

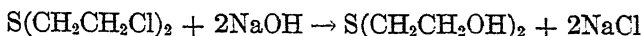
A simpler but less obviously rational form of this equation is:

$$dC/dt = 2k_1(M)/[1 + k_3(\text{Cl}^-)] = 0.304(M)/[1 + 6.14(\text{Cl}^-)]$$

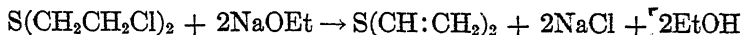
This equation has been found to hold with satisfactory accuracy for a wide variety of acid solutions.

Helfrich and Reid (29) state that when bis( $\beta$ -chloroethyl) sulfide is acted on with 20 per cent sodium ethoxide, the sodium hydroxide present

hydrolyzes the sulfide with separation of sodium chloride from the alcohol solution, the reaction going for the most part according to the equation



However, another reaction takes place to some extent, forming an easily polymerizable liquid, which they surmised to be divinyl sulfide, thus:



This was later confirmed by Bales and Nickelson (5).

Bales and Nickelson (6) increased the 26 per cent yield of divinyl sulfide by distilling bis( $\beta$ -chloroethyl) sulfide with 20 per cent potassium hydroxide without previous heating under reflux. Hydrolysis with alcoholic potas-

TABLE 14

*Hydrolysis of bis( $\beta$ -chloroethyl) sulfide with 20 per cent potassium hydroxide in alcohol-water mixtures*

WATER	YIELD	DENSITY
<i>per cent</i>	<i>per cent*</i>	
5	35.0	0.931
15	33.5	0.941
17	33.6	0.960
20	33.5	1.014
24	21.0	1.034
43	16.0	1.050

\* The percentage of the product of hydrolysis calculated on the weight of bis( $\beta$ -chloroethyl) sulfide taken.

sium hydroxide containing water causes the product to increase in density as the proportion of water increases, until, when the alcohol contains 20 per cent water, the hydrolysis product has a density greater than that of water. The yield is almost constant up to this concentration of water, but with higher concentration the yields are smaller and the product more dense as shown in table 14. With alcoholic sodium ethoxide, the yield was 36 per cent.

When the product obtained by the hydrolysis of bis( $\beta$ -chloroethyl) sulfide with 20 per cent potassium hydroxide in 80:20 or 57:43 alcohol-water was heated and carefully purified, a clear, colorless liquid boiling at 152–154°C.,  $d_4^{15} = 1.110$ , was obtained. This liquid, which contained chlorine and which regenerated bis( $\beta$ -chloroethyl) sulfide on treatment with phosphorus trichloride, was thought to be  $\beta$ -chloro- $\beta'$ -hydroxydiethyl sulfide,  $\text{ClCH}_2\text{CH}_2\cdot\text{S}\cdot\text{CH}_2\text{CH}_2\text{OH}$ ; this was later confirmed by Peters and Walker (57).



Boulin and Simon (9) found that when 1 part of the sulfide was dissolved in 200 parts of water, the ratio of the length of time for complete hydrolysis at 20°, 50°, and 100°C. is 1:7.5:30. A longer time is required if less water is used. The sulfide decomposes and distills with the water vapor at 100°C. Acids and ammonium hydroxide (1 mole per liter) slightly retard the decomposition; 100 cc. of water containing 100 millimoles of potassium hydroxide per liter decomposed 7.5 millimoles of the sulfide in 50 hours at 50°C.; 3 hours suffice in the absence of the hydroxide, the retardation being due to the potassium chloride formed.

Kramer (35) has shown that when bis( $\beta$ -chloroethyl) sulfide is shaken with a 1 per cent solution of sodium silicate, after the elapse of two hours each drop of the sulfide is surrounded by a sheath of colloidal silicic acid, and in 48 hours the whole tube is filled with colloidal silicic acid gel. The sulfide is hydrolyzed to dihydroxydiethyl sulfide, and the hydrogen chloride set free liberates colloidal silicic acid.

Davies and Oxford (17) hydrolyzed bis( $\beta$ -chloroethyl) sulfide with 4 moles of 20 per cent potassium hydroxide in alcohol-water and obtained therefrom several compounds: divinyl sulfide,  $S(CH:CH_2)_2$ , boiling at 85°C., a colorless, highly refractive, mobile oil;  $\beta$ -ethoxyethyl vinyl sulfide,  $CH_2:CH \cdot S \cdot CH_2CH_2OEt$ , a colorless mobile liquid with a curious, pungent, camphor-like odor, boiling at 65°C. at 8 mm.,  $d_4^{20} = 0.9532$ ;  $\beta$ -ethoxy- $\beta'$ -hydroxydiethyl sulfide,  $HOCH_2CH_2 \cdot S \cdot CH_2CH_2OEt$ , boiling at 117.5°C. at 4 mm. pressure; bis( $\beta$ -ethoxyethyl) sulfide,  $S(CH_2CH_2OEt)_2$ , boiling at 101–102°C. at 4 mm.,  $d_4^{20} = 0.9658$ . Using 1 mole of 20 per cent potassium hydroxide in alcohol-water,  $\beta$ -chloroethyl vinyl sulfide,  $CH_2:CH \cdot S \cdot CH_2CH_2Cl$ , boiling at 71–72°C. at 50 mm. and at 151.5–152.5°C. at 764 mm., is obtained.

Wilson, Fuller, and Schur (68) determined the acceleration of the hydrolysis of bis( $\beta$ -chloroethyl) sulfide by alkaline colloidal solutions. With a solution containing 3 per cent sulfonated corn oil and 2 per cent sodium carbonate the hydrolysis of bis( $\beta$ -chloroethyl) sulfide on cloth is completed in 35 minutes at 85°C. The alkaline sulfonated corn oil accelerates only the specific rate of solution per unit area of interface, by acting as carrier to transport the sulfide from the very thin aqueous layer through which the gas was diffusing before hydrolysis was complete; it has no tendency to increase the specific rate of hydrolysis nor the solubility of the gas in the aqueous phase. Alkaline soap solutions do not complete the hydrolysis of the sulfide on cloth, because of the formation of a gummy compound, probably a fatty acid ester of thiodiglycol, between the hydrolysis product of the gas and soap, which deposits on the films of the cloth.

### 3. Oxidation

Bis( $\beta$ -chloroethyl) sulfide is ultimately oxidized by various reagents to the corresponding sulfone:



When pure bis( $\beta$ -chloroethyl) sulfide is dropped into concentrated nitric acid at room temperature, there is an energetic reaction with evolution of heat and formation of a light green solution (21, 29, 44). From this solution, after diluting with water, the sulfoxide separates as white crystalline plates which melt at 109.5–110°C.; so far as can be judged by handling, it is entirely without effect upon the human skin. The sulfoxide is soluble in water, concentrated mineral acids, alcohol, ether, benzene, carbon disulfide, and acetone. It can not be distilled without partial decomposition even at reduced pressure, the chief decomposition product being the sulfide; it seems that during the distillation the sulfoxide dissociates into the sulfide and oxygen, the former appearing in the distillate and the latter escaping as such or effecting a destructive oxidation of the sulfoxide. Upon brominating an ice-cold chloroform solution of the sulfide a deep orange-colored precipitate quickly separates which, after drying and purifying, results in a bright yellow solid which yields the sulfoxide upon treatment with dilute sodium hydroxide solution (21). The sulfoxide is also prepared by slowly treating the sulfide in cold acetic acid with 30 per cent hydrogen peroxide and allowing to stand overnight (60).

When the sulfoxide is further oxidized by the use of fuming nitric acid, potassium permanganate, or chromic acid, the chief product is the corresponding sulfone which, when recrystallized from water or alcohol, appears as beautiful thin white leaflets melting at 48°C. (60), 53.5°C. (16), 56°C.; it distills *in vacuo* without decomposition at 183°C. at 20 mm. (29). Its solubility in water and dilute acids is much less than that of the sulfoxide. Bennett (7) obtained  $\beta$ -chloroethanesulfonic acid,  $\text{ClCH}_2\text{CH}_2\text{SO}_3\text{H}$ , when he treated the sulfoxide with fuming nitric acid in a sealed tube for 10 hours at 100°C.

The sulfone produces, on contact with the skin, blisters and persistent sores; its vapors are lachrymatory, causing a copious flow of tears and a soreness of the eyes, and are also somewhat sternutatory, causing sneezing and a slight burning in the nostrils, but as its vapor pressure is low at ordinary temperatures no inconvenience is experienced in handling it. That these unpleasant properties should reoccur in the sulfone after disappearing in the sulfoxide is peculiar. The theory has been advanced (48)

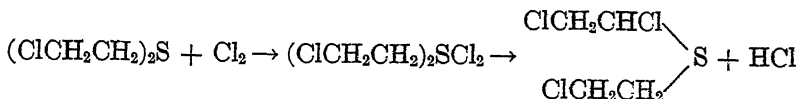
that hydrolysis is one of the main factors entering into the action of the sulfide on the human body. It is of interest to note that hydrolysis of the sulfide by water is greatest, the sulfone next, and the sulfoxide least. These factors are in accord with the theory as applied to the observed physiological effects of the sulfoxide and sulfone.

#### 4. Chlorination

Chlorine is slowly but steadily absorbed when passed through bis( $\beta$ -chloroethyl) sulfide; after a short time hydrogen chloride is evolved and the temperature rises to about 50°C. The product boils with partial decomposition under atmospheric pressure, and when distilled under reduced pressure yields no constant-boiling compound. On distillation under 15 mm. through a 65 mm. Dufton column at the rate of one drop each three seconds, Mann and Pope (44) resolved the mixtures, and after several similar fractional distillations the following three compounds were obtained in a state of purity: (a)  $\alpha, \beta, \beta'$ -trichlorodiethyl sulfide,  $\text{ClCH}_2\text{-CH}_2\text{-S-CHClCH}_2\text{Cl}$ , boiling at 106.5–108°C. at 15 mm.,  $d_4^{21.8^\circ} = 1.4219$ ; (b)  $\alpha, \beta, \beta, \beta'$ -tetrachlorodiethyl sulfide,  $\text{ClCH}_2\text{CH}_2\text{-S-CHClCHCl}_2$ , boiling at 123–125°C. at 15 mm.,  $d_4^{15.2^\circ} = 1.5441$ ; (c)  $\alpha, \alpha, \beta, \beta, \beta, \beta'$ -hexachlorodiethyl sulfide,  $\text{ClCH}_2\text{CH}_2\text{-S-CCl}_2\text{CCl}_2$ , boiling at 160–161°C. at 15 mm.,  $d_4^{10.6^\circ} = 1.6944$ . These three chlorinated diethyl sulfides are colorless liquids which slowly acquire a light green tint when exposed to sunlight; they possess a faint odor similar to the parent dichloro compound, but are devoid of vesicant properties and have much lower melting points. As the number of chlorine atoms in the molecule increases, the sulfur atom shows increasing reluctance to become quadrivalent.

On passing dry chlorine into bis( $\beta$ -chloroethyl) sulfide, allowing the liquid to warm up spontaneously, and finally heating at 100°C. in the current of chlorine, sulfur dichloride distils over, the absorption ceasing after about twenty-two hours; the product, on distillation, yields a mixture of sulfur monochloride and dichloride followed by an oil boiling at 130–175°C. Above this temperature the distillate solidifies in the condenser, the crystalline substance being identified as perchloroethane, melting at 185°C. The liquid boiling at 130–175°C. separates sharply, on distillation through a Dufton column, into  $\alpha, \beta, \beta, \beta$ -tetrachloroethane, boiling at 129–130°C., and pentachloroethane, boiling at 159–161°C., while a crystalline residue of perchloroethane remains in the distilling flask.

Because of the ease and rapidity with which bis( $\beta$ -chloroethyl) sulfide is chlorinated, Lawson and Dawson (37) suggest that a preliminary addition of the chlorine followed by a molecular rearrangement involving substitution of one atom of chlorine and elimination of one molecule of hydrogen chloride takes place as follows:



Mumford and Phillips (51) advanced the view that the products of chlorination of bis( $\beta$ -chloroethyl) sulfide described by Mann and Pope (44) as  $\alpha,\beta,\beta'$ -trichloro- and  $\alpha,\beta,\beta,\beta'$ -tetrachlorodiethyl sulfides were  $\alpha,\beta$ -dichloro- and  $\alpha,\beta,\beta$ -trichlorovinyl  $\beta$ -chloroethyl sulfide respectively.

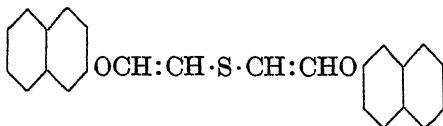
Bis( $\beta$ -chloroethyl) sulfide in carbon tetrachloride and chlorine at  $-5^\circ\text{C}$ . to  $0^\circ\text{C}$ . gives the chlorine intermediate addition product in the process of chlorination, bis( $\beta$ -chloroethyl) sulfonium chloride,  $(\text{ClCH}_2\text{CH}_2)_2\text{SCl}_2$ , which Lawson and Dawson (37) crystallized as white needles and which was found to decompose at room temperature to give hydrogen chloride,  $\alpha,\beta,\beta'$ -trichlorodiethyl sulfide, and dichloroethyl sulfoxide. Distillation of  $\alpha,\beta,\beta'$ -trichlorodiethyl sulfide gave two isomeric vinyl compounds;  $\beta$ -chloroethyl  $\beta$ -chlorovinyl sulfide and  $\beta$ -chloroethyl  $\alpha$ -chlorovinyl sulfide. These investigators found that on distillation of this relatively pure trichlorodiethyl sulfide, great quantities of hydrogen chloride were evolved but no definite boiling point was observed. After ten distillations two distinct fractions were obtained which analysis showed to have the same empirical formula corresponding to chloroethyl chlorovinyl sulfide.

The instability of  $\alpha,\beta,\beta'$ -trichlorodiethyl sulfide and its decomposition on distillation to give two isomers affords an explanation for the failure of Mann and Pope to isolate the pentachlorodiethyl sulfide. They found the trichlorodiethyl sulfide so unstable that it was impossible to isolate it by distillation at pressures as low as 3.5 mm.; the compound isolated by Mann and Pope could not, therefore, have been the trichlorodiethyl sulfide. They postulated the splitting-out of hydrogen chloride from tetrachlorodiethyl sulfide to give a chloroethyl dichlorovinyl sulfide, and similarly the decomposition of pentachlorodiethyl sulfide to give chloroethyl vinyl sulfide, which was afterwards confirmed by Phillips, Davies, and Mumford (58). An explanation was thus afforded for the inability of Mann and Pope to isolate the pentachloro derivative.

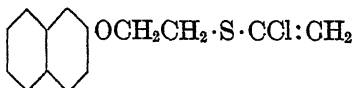
Contrary to previous views Phillips, Davies, and Mumford (58) proved that chlorination is not confined to one chain but occurs in both chloroethyl groups.

Besides compounds heretofore mentioned the following bis( $\beta$ -chloroethyl) sulfide chlorinated products have been prepared and identified (37):  $\alpha,\beta,\beta'$ -trichlorodiethyl sulfide,  $\text{CHCl}_2\text{CHCl} \cdot \text{S} \cdot \text{CH}_2\text{CH}_2\text{Cl}$ , a colorless liquid, decomposing at room temperature with evolution of hydrogen chloride,  $d_4^{20} = 1.4038$ ,  $n_D^{22.5} = 1.5309$ ;  $\beta$ -chloroethyl  $\beta$ -chlorovinyl sulfide,  $\text{CHCl} : \text{CH} \cdot \text{S} \cdot \text{CH}_2\text{CH}_2\text{Cl}$ , a colorless oil, boiling at  $73.5\text{--}74.5^\circ\text{C}$ . at 4 mm.,

$d_4^{20} = 1.3280$ ,  $n_D^{22.5} = 1.5483$ ;  $\beta$ -chloroethyl  $\alpha$ -chlorovinyl sulfide,  $\text{CH}_2\text{Cl}-\text{CH}_2\cdot\text{S}\cdot\text{CH}_2\cdot\text{CCl}$ , a colorless liquid, boiling at  $69.5\text{--}70.5^\circ\text{C}$ . at 4 mm.,  $d_4^{20} = 1.3193$ ,  $n_D^{22.5} = 1.550$ ;  $\alpha$ -bromo- $\beta$ ,  $\beta'$ -dichloroethyl sulfide,  $\text{CH}_2\text{Cl}-\text{CHBr}\cdot\text{S}\cdot\text{CH}_2\text{CH}_2\text{Cl}$ , a yellow oil, unstable at room temperature; bis( $\beta$ -naphthoxy)ethyl vinyl sulfide,



tiny diamond-shaped plates, melting at  $129.5^\circ\text{C}$ .;  $\beta$ -naphthoxyethyl  $\alpha$ -chlorovinyl sulfide,



an oil which crystallizes from petroleum ether giving crystals melting at  $56\text{--}57^\circ\text{C}$ .;  $\alpha$ ,  $\beta$ ,  $\beta'$ -tetrachlorodiethyl sulfide,  $\text{CHCl}_2\text{CHCl}\cdot\text{S}\cdot\text{CH}_2\text{CH}_2\text{Cl}$ , boiling at  $118.5\text{--}119^\circ\text{C}$ . at 3 mm.;  $\alpha$ ,  $\alpha$ ,  $\beta$ ,  $\beta'$ -tetrachlorodiethyl sulfide,  $\text{CH}_2\text{ClCCl}_2\cdot\text{S}\cdot\text{CH}_2\text{CH}_2\text{Cl}$ , too unstable for isolation;  $\beta$ -chloroethyl  $\alpha$ ,  $\beta$ -dichlorovinyl sulfide,  $\text{CH}_2\text{ClCH}_2\cdot\text{S}\cdot\text{CCl}:\text{CHCl}$ , a colorless mobile liquid, boiling at  $79.7\text{--}80.5^\circ\text{C}$ . at 4 mm.,  $d_4^{20} = 1.4315$ ,  $\eta^{20} = 0.028$  dynes per square centimeter (58);  $\beta$ -chloroethyl  $\beta$ ,  $\beta$ -dichlorovinyl sulfide,  $\text{CH}_2\text{Cl}-\text{CH}_2\cdot\text{S}\cdot\text{CH}:\text{CCl}_2$ , boiling at  $90\text{--}91^\circ\text{C}$ . at 3.5 mm.;  $\beta$ -chloroethyl  $\alpha$ ,  $\beta$ ,  $\beta$ -trichlorovinyl sulfide,  $\text{CH}_2\text{ClCH}_2\cdot\text{S}\cdot\text{CCl}:\text{CCl}_2$ , boiling at  $97\text{--}98^\circ\text{C}$ . at 4 mm., and at  $123.5\text{--}124.5^\circ\text{C}$ . at 15 mm.;  $\alpha$ ,  $\alpha$ ,  $\beta$ ,  $\beta$ ,  $\beta'$ -pentachlorodiethyl sulfide,  $\text{CHCl}_2\text{CCl}_2\cdot\text{S}\cdot\text{CH}_2\text{CH}_2\text{Cl}$ , an unstable yellow oil;  $\alpha$ ,  $\alpha$ ,  $\beta$ ,  $\beta$ ,  $\beta'$ -hexachlorodiethyl sulfide,  $\text{CCl}_3\text{CCl}_2\cdot\text{S}\cdot\text{CH}_2\text{CH}_2\text{Cl}$ , boiling at  $131\text{--}132^\circ\text{C}$ . at 4.5 mm., and at  $159.5\text{--}60^\circ\text{C}$ . at 15 mm.,  $d_4^{20} = 1.6849$ ,  $\eta^{20} = 0.254$  dynes per square centimeter (58).  $\beta$ ( $\beta$ -Chloroethylthiol)ethyl trichlorovinyl

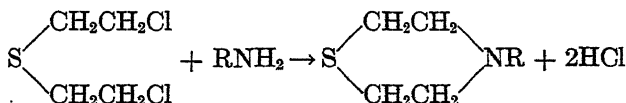
sulfide,  $\text{S}$  , crystallizes from alcohol in long flat needles,

boils at  $157\text{--}167^\circ\text{C}$ . at 2.5 mm., melts at  $70.5^\circ\text{C}$ .;  $\alpha$ ,  $\alpha$ ,  $\beta$ -trichloroethyl  $\beta$ -chlorovinyl sulfide,  $\text{CH}_2\text{ClCCl}_2\cdot\text{S}\cdot\text{CH}:\text{CHCl}$ , a colorless mobile liquid, boiling at  $122\text{--}123^\circ\text{C}$ . at 15 mm.,  $d_4^{20} = 1.5404$ ,  $\eta^{20} = 0.057$  dynes per square centimeter;  $\alpha$ ,  $\beta$ -dichloroethyl  $\alpha$ ,  $\beta$ -dichlorovinyl sulfide,  $\text{CHCl}:\text{CCl}\cdot\text{S}\cdot\text{CHClCH}_2\text{Cl}$  a colorless liquid, boiling at  $120\text{--}121^\circ\text{C}$ . at 15 mm.,  $d_4^{15.2} = 1.544$ ;  $\alpha$ ,  $\alpha$ ,  $\alpha'$ ,  $\beta'$ ,  $\beta'$ -hexadichlorodiethyl sulfide,  $\text{CH}_2\text{ClCCl}_2\cdot\text{S}\cdot\text{CHCl}-\text{CHCl}_2$ , a colorless liquid, boiling at  $159\text{--}160^\circ\text{C}$ . at 15 mm.,  $d_4^{20} = 1.6841$ ,  $\eta^{20} = 0.294$  dynes per square centimeter;  $\alpha$ ,  $\beta$ ,  $\beta$ -trichloroethyl  $\alpha$ ,  $\beta$ -

dichlorovinyl sulfide,  $\text{CHCl}:\text{CCl}\cdot\text{S}\cdot\text{CHClCHCl}_2$ , a colorless liquid, boiling at  $134\text{--}135^\circ\text{C}$ . at 15 mm.,  $d_4^{20} = 1.6293$ ,  $\eta^{20} = 0.065$  dynes per square centimeter;  $\alpha,\alpha,\alpha',\beta,\beta,\beta',\beta'$ -heptachlorodiethyl sulfide,  $\text{CHCl}_2\text{CCl}_2\cdot\text{S}\cdot\text{CHClCHCl}_2$ , a colorless viscid liquid, boiling at  $170\text{--}172^\circ\text{C}$ . at 15 mm.,  $d_4^{20} = 1.7473$ ,  $\eta^{20} = 0.427$  dynes per square centimeter;  $\alpha,\alpha,\alpha',\beta,\beta,\beta'$ -hexachlorodiethyl sulfide,  $\text{CHCl}_2\text{CCl}_2\cdot\text{S}\cdot\text{CHClCH}_2\text{Cl}$ , boiling at  $157\text{--}159^\circ\text{C}$ . at 15 mm.,  $d_4^{20} = 1.6825$ ,  $\eta^{20} = 0.273$  dynes per square centimeter;  $\alpha,\beta$ -dichloroethyl  $\alpha,\beta$ -trichlorovinyl sulfide,  $\text{CCl}_2:\text{CCl}\cdot\text{S}\cdot\text{CHClCH}_2\text{Cl}$ , boiling at  $133\text{--}134^\circ\text{C}$ . at 15 mm.,  $d_4^{20} = 1.6190$ ,  $\eta^{20} = 0.063$  dynes per square centimeter;  $\alpha,\alpha,\alpha',\beta,\beta,\beta,\beta'$ -heptachlorodiethyl sulfide,  $\text{CCl}_3\text{CCl}_2\cdot\text{S}\cdot\text{CHClCH}_2\text{Cl}$ , boiling at  $132\text{--}134^\circ\text{C}$ . at 2 mm.,  $d_4^{20} = 1.7430$ ,  $\eta^{20} = 0.382$  dynes per square centimeter;  $\alpha,\alpha,\alpha,\beta$ -tetrachloroethane,  $\text{CCl}_3\text{CH}_2\text{Cl}$ , b.p.  $35\text{--}36^\circ\text{C}$ . at 20 mm.,  $129\text{--}129.7^\circ\text{C}$ . at 760 mm.;  $\alpha,\beta$ -dichloroethyl  $\alpha,\beta,\beta'$ -trichlorovinyl sulfide,  $\text{CCl}_2:\text{CCl}\cdot\text{S}\cdot\text{CHClCH}_2\text{Cl}$ , boiling at  $133.5\text{--}134^\circ$  at 15 mm.,  $d_4^{20} = 1.6183$ ,  $1.6183$ ,  $\eta^{20} = 0.069$  dynes per square centimeter.

#### 4. Formation of primary amines

Primary amines, thiazanes, are prepared from bis( $\beta$ -chloroethyl) sulfide and the amine in ethyl alcohol containing sodium carbonate.

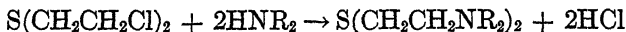


They are colorless, mobile oils with densities less than 1 (except the benzyl derivative).

The following thiazanes have been prepared: *methyl* (13), completely miscible with water and organic solvents, a powerful base fuming in the presence of acid vapors, b.p.  $163\text{--}164^\circ\text{C}$ . at 757 mm.,  $d_4^{15} = 0.99596$ ; *ethyl* (13), insoluble in water, b.p.  $184^\circ\text{C}$ . at 763 mm.,  $d_4^{15} = 0.99295$ ; *isoamyl* (13), oil with characteristic odor, almost insoluble in water, b.p.  $105\text{--}106^\circ\text{C}$ . at 13 mm.; *benzyl* (13), insoluble in water but soluble in organic solvents, b.p.  $154^\circ\text{C}$ . at 13 mm.,  $152.5^\circ\text{C}$ . at 10 mm. (39),  $d_4^{15} = 1.08328$ . *Propyl*, soluble in water, b.p.  $82.5\text{--}83.5^\circ\text{C}$ . at 12 mm.; *isopropyl*, soluble in water, b.p.  $74.5\text{--}75^\circ\text{C}$ . at 9 mm.; *butyl*, b.p.  $118^\circ\text{C}$ . at 33 mm.,  $101\text{--}102^\circ\text{C}$ . at 19 mm.; *isobutyl*, b.p.  $93^\circ\text{C}$ . at 17 mm.; *amyl*, b.p.  $119^\circ\text{C}$ . at 15 mm.,  $104^\circ\text{C}$ . at 9 mm.,  $d_4^{15} = 0.9485$  (39).

#### 5. Formation of secondary amines

These are prepared in the same manner as the corresponding primary amines except that two molecular equivalents of secondary amine are used for each equivalent of sulfide.



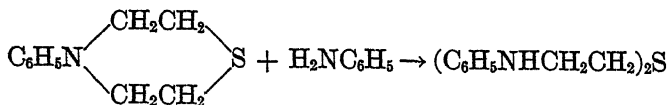
All the bis( $\beta$ -dialkylaminoethyl) sulfides are colorless oils, soluble in organic solvents but almost insoluble in water except for the diethyl derivative, which has a solubility of about 0.1 per cent (39). The following have been prepared: *Dimethylaminoethyl vinyl sulfide*, one amino group was lost thus forming the unsaturated compound, b.p. 168.5°C. at 763 mm.; *ethyl*, b.p. 139–140°C. at 9 mm.; *propyl*, b.p. 194°C. at 19 mm.,  $d_4^{25} = 0.9007$ ; *butyl*, b.p. 205–206°C. at 10 mm.; *piperidyl*, b.p. 174.5–176°C. at 15 mm.; *phthalimide*, m.p. 125°C., fine, pale cream-colored needles (11).

#### 6. Formation of tertiary amines

Condensation with tertiary amines is by addition, unaccompanied by the elimination of any by-product. When trimethylamine and bis( $\beta$ -chloroethyl) sulfide are heated in a sealed tube for 6 hours at 100°C., crystals of  $\text{S}(\text{CH}_2\text{CH}_2\text{ClNMe}_3)_2$  are formed, which melt at 135.5°C. (39).

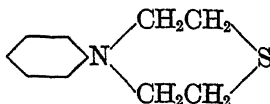
#### 7. Formation of aromatic amines

The aromatic derivatives are prepared with difficulty, this difficulty being based on two facts: (1) their insolubility (which causes inconvenience in purification), and (2) their secondary reaction with excess aromatic amines, apparently to break open the heterocyclic ring:



The compound resulting from the secondary reaction is easily soluble in dilute hydrochloric acid (quite in contrast with the thiazanes), yielding in alkaline solution an oil heavier than water, which is only slightly volatile in steam.

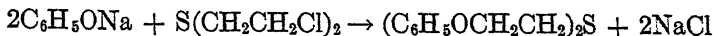
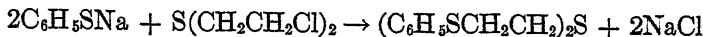
4-Phenyl-1-thiazane,



a dull white powder, not pure, soluble in hot toluene, m.p. 108–111°C.

#### 8. Condensation with sodium thiophenate and phenates

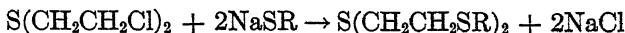
The fundamental reaction is the formation of sodium chloride with the simultaneous condensation of the organic residues:



Thus it is seen that the condensation with sodium thiophenate gives complex trisulfides or related compounds, while with the phenates it produces various mixed sulfide-ethers. All that have been prepared by Helfrich and Reid (29) are white, or nearly white, crystalline bodies and most of them can be recrystallized from alcohol but are insoluble in water (see table 15).

### 9. Condensation with sodium mercaptides (29)

The condensation of bis( $\beta$ -chloroethyl) sulfide with sodium mercaptides is quite similar to that with sodium thiophenate already considered:



The reaction is rapid in the hot alcohol solution and the yield nearly quantitative. The stability of these multiple sulfides and their oxidation products is again an outstanding feature. Data concerning them are given in table 16.

### 10. Miscellaneous reactions

Upon distilling at reduced pressure Mann and Pope (44) obtained a trichlorodiethyl sulfide and tetrachlorodiethyl sulfide from a mixture of bis( $\beta$ -chloroethyl) sulfide and sulfur monochloride; it thus appears that sulfur monochloride acts as a chlorinating agent on bis( $\beta$ -chloroethyl) sulfide.

Upon heating bis( $\beta$ -chloroethyl) sulfide under pressure with alcoholic ammonia Clarke (13) obtained 1,4-thiazane,  $\text{CH}_2 \cdot \text{S} \cdot \text{CH}_2\text{CH}_2\text{NHCH}_2$ , a colorless liquid, boiling at  $169^\circ\text{C}$ . at 758 mm., miscible with water and organic solvents, possessing an odor similar to, yet more unpleasant than, that of pyridine. It is a very strong base, very rapidly absorbing carbon dioxide from the air, and reacts vigorously with metallic potassium on gentle warming.

Bis- $\beta$ -thioethyl acetate,  $\text{S}(\text{CH}_2\text{CH}_2\text{OOCMe})_2$ , is prepared (29) by distilling *in vacuo* the oil obtained after anhydrous sodium acetate and bis( $\beta$ -chloroethyl) sulfide have been heated; b.p.  $155\text{--}156^\circ\text{C}$ . at 20 mm.,  $d^{20} = 1.132$ ,  $n^{22.5} = 1.4720$ .

$\beta, \beta'$ -Dichlorodiethylsulfine-*p*-toluenesulfonylimine,

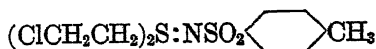





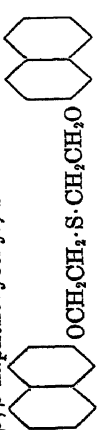
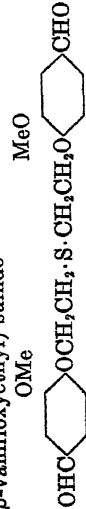
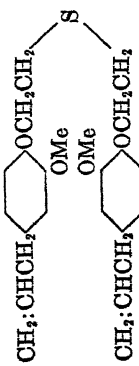
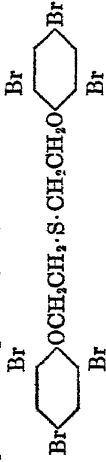




TABLE 15  
*Products formed by condensation of bis( $\beta$ -chloroethyl) sulfide with sodium thiophenate and phenates*

NAME AND FORMULA	MELTING POINT °C.	SOLUBILITY IN 100 CC. ALCOHOL		REMARKS
		Cold (18°C.) grams	Hot (75°C.) grams	
Bis( $\beta$ -phenylmercaptoethyl) sulfide (C <sub>6</sub> H <sub>5</sub> SC <sub>2</sub> H <sub>4</sub> ) <sub>2</sub> S	57.5	0.65	10.00	Lustrous white leaflets
Bis( $\beta$ -phenoxyethyl) sulfide (C <sub>6</sub> H <sub>5</sub> OCH <sub>2</sub> CH <sub>2</sub> ) <sub>2</sub> S	54.2	1.42	20.00	Fine white needles
Bis( $\beta$ - <i>p</i> -cresoxyethyl) sulfide (CH <sub>3</sub>  OCH <sub>2</sub> CH <sub>2</sub> ) <sub>2</sub> S	78.0	0.20	Soluble	Lustrous needles
Bis( $\beta$ - <i>o</i> -cresoxyethyl) sulfide  OCH <sub>2</sub> CH <sub>2</sub> -S-CH <sub>2</sub> CH <sub>2</sub>  Me	46.5	0.30	Soluble	White needles
Bis( $\beta$ , $\alpha$ -naphthoxyethyl) sulfide  OCH <sub>2</sub> CH <sub>2</sub> -S-CH <sub>2</sub> CH <sub>2</sub> O 	94.5	Insoluble	Slightly soluble	Light brown crystals

<p>Bis(<math>\beta, \beta</math>-naphthoxyethyl) sulfide</p> 	129.0	Insoluble	0.60	Small crystals
<p>Bis(<math>\beta</math>-vanilloxyethyl) sulfide</p> 	131.5	Insoluble	2.00	Gray needles
<p>Bis(<math>\beta</math>-eugenoxoethyl) sulfide</p> 	113.5	0.10	2.00	Yellow powder
<p>Bis(<math>\beta</math>-tribromophenoxyethyl) sulfide</p> 	118.5	Insoluble	0.50	White powder

is obtained (45) in the form of fine white needles, m.p. 144–145°C., upon shaking bis( $\beta$ -chloroethyl) sulfide with an aqueous solution of chloramine-T.

Cashmore and McCombie (11) prepared the ethyl ester of bis( $\beta$ -chloroethyl) sulfide,  $(\text{CH}_2\text{CH}_2\text{NHCH}_2\text{CO}_2\text{Et})_2\text{S}$ , from bis( $\beta$ -chloroethyl) sulfide and glycine ester hydrochloride in 98 per cent ethyl alcohol containing anhydrous sodium acetate. It is an oil boiling at 159–160°C. at 15 mm.

By boiling for several hours a mixture of potassium cyanide and bis( $\beta$ -chloroethyl) sulfide the nitrile, probably  $(\text{CH}_2\text{SCH}_2\text{CH}_2\text{CN})_2$ , is obtained (16). It crystallizes in colorless microneedles from benzene, melting at 91–91.5°C., and readily soluble in alcohol but insoluble in water.

TABLE 16

*Compounds formed by the condensation of bis( $\beta$ -chloroethyl) sulfide with sodium mercaptides*

NAME AND FORMULA	BOILING POINT	MELT- ING POINT	REMARKS
	<i>degrees C.</i>	<i>degrees C.</i>	
Bis( $\beta$ -ethylmercaptoethyl) sulfide $(\text{EtSCH}_2\text{CH}_2)_2\text{S}$	173–175 at 21 mm.	17.0 *	$n^{22.5^\circ} = 1.5151$ ; odor rather unpleasant and persistent
Bis( $\beta$ -propylmercaptoethyl) sulfide $(\text{PrSCH}_2\text{CH}_2)_2\text{S}$	193–195 at 22 mm.	27.5	
Bis( $\beta$ -butylmercaptoethyl) sulfide $(\text{BuSCH}_2\text{CH}_2)_2\text{S}$	222–223 at 21 mm.	17.5	Oil; unpleasant, per- sistent odor; $d^{20^\circ} =$ 1.102; $n^{22.5^\circ} =$ 1.5455
Bis( $\beta$ -amylmercaptoethyl) sulfide $(\text{isoAmSCH}_2\text{CH}_2)_2\text{S}$	170 at 4 mm.	20.0	

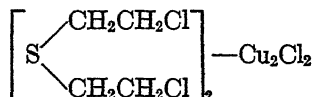
$\gamma$ -Thiodibutyric acid,  $\text{S}(\text{CH}_2\text{CH}_2\text{CH}_2\text{COOH})_2$ , is formed when the product of the condensation of the sulfide with ethyl malonate (or ethyl acetoacetate) and sodium ethoxide is subjected to alkaline hydrolysis (16). The acid, obtained as colorless leaves with a nacreous luster, m.p. 97–98°C., is fairly soluble in hot water, moderately so in alcohol, ether, and benzene, and does not distil in a current of steam or under diminished pressure.

Diiodoethyl sulfide,  $(\text{CH}_2\text{CH}_2\text{I})_2\text{S}$ , results from the condensation of bis( $\beta$ -chloroethyl) sulfide with an alcoholic solution of sodium iodide (29). It is a yellowish white crystalline material, m.p. 55–60°C., darkens slowly at room temperature, is insoluble in water and only slightly soluble in cold and moderately soluble in hot alcohol.

When bis( $\beta$ -chloroethyl) sulfide is heated several hours on a water bath with a 20 per cent solution of sodium thiocyanate there is formed bis( $\beta$ -thiocyanoethyl sulfide,  $S(CH_2CH_2SCN)_2$ , which crystallizes as light yellow crystals from ether, m.p.  $35^\circ C$ . (33).

#### IV. DETECTION AND DETERMINATION

The usual procedure for the estimation of bis( $\beta$ -chloroethyl) sulfide in a sample is to distil a known volume under diminished pressure and to collect the fraction boiling at  $125$ – $130^\circ C$ . at 40 mm. This product is fairly pure, and the volume obtained affords an approximation of the purity of the crude product. This procedure involves personal risks, and gives no accurate figure for the absolute percentage of bis( $\beta$ -chloroethyl) sulfide, as the latter fraction also contains higher chlorinated compounds and other impurities. With the object of overcoming these difficulties Hollely (32) devised the following volumetric method for the estimation of bis( $\beta$ -chloroethyl) sulfide with or without solvents. This method depends upon the fact that the sulfide forms a definite double salt with cuprous chloride, of the constitution



As the higher chlorinated compounds do not react with cuprous chloride, the method is applicable for the absolute estimation of the sulfide.

When a 10 per cent solution of bis( $\beta$ -chloroethyl) sulfide in glacial acetic acid is heated with a 67 per cent solution of hydriodic acid at  $50$ – $70^\circ C$ ., a yellowish oil is deposited which solidifies on cooling to a crystalline mass of diiodoethyl sulfide, melting at  $62^\circ C$ . Grignard, Rivat, and Scatchard (23) make use of this in the determination of bis( $\beta$ -chloroethyl) sulfide. About 1 g. (weight =  $P$ ) of sample is introduced into a small flask fitted with an air condenser and containing 15 g. of glacial acetic acid and exactly 5 cc. of 54 per cent hydriodic acid; the whole is heated 15 minutes at  $70^\circ C$ . After cooling, the contents of the flask is washed into a 500-cc. volumetric flask containing exactly 100 cc. of carbon tetrachloride and about 200 cc. of water. After dissolving the precipitate, the flask is filled to the mark with water, shaken well, and 50 cc. of the aqueous layer treated with 10 cc. of 10 per cent sodium nitrite. The liberated iodine is extracted with carbon tetrachloride which is then washed with 100 cc. of water, and the iodine titrated with 0.1  $N$  sodium thiosulfate, the volume used being  $A_1$  cc. The carbon tetrachloride solution in the 500-cc. flask is then separated carefully, the aqueous solution washed with a little carbon tetrachloride, which is added to the main solution, and the combined carbon tetrachloride

extracts are titrated with 0.1 *N* sodium thiosulfate, the volume consumed being  $A_2$  cc. A blank determination of the total iodine in 5 cc. of the hydriodic acid solution used is made, the 0.1 *N* sodium thiosulfate consumed being  $A_0$  cc. The percentage of bis( $\beta$ -chloroethyl) sulfide in the sample is then  $0.82/P [10A_0 + 1.5 - (8A_1 + A_2)]$ .

Yablick, Perrott, and Furman (69) determined quantitatively the presence of traces of bis( $\beta$ -chloroethyl) sulfide in the air by reducing a 1 per cent solution of selenious acid in dilute sulfuric acid by means of the sulfide vapor to an orange-red suspension of selenium. Amounts of the sulfide as small as 0.005 mg. can be detected in this way. The amount of suspension produced is nearly directly proportional to the amount of the sulfide present, and 0.10 to 0.01 mg. can be established nephelometrically with a fair degree of accuracy ( $\pm 0.005$  mg.). This method is remarkably free from the necessity for extreme purity of original solutions, cleanliness of glassware, etc., but the selenious acid reagent is *not* specific to bis( $\beta$ -chloroethyl) sulfide, for arsine and the substituted arsines and several other toxic gases react with it vigorously.

Spica (59) finds nitrosyl iodide to be the *only* specific reagent proposed for detection of bis( $\beta$ -chloroethyl) sulfide which gives the corresponding diiodo compound.

Maxim (49) oxidizes the sulfur of the sulfide by heating with potassium permanganate, then passing the gas into barium chloride solution containing hydrogen peroxide. The barium sulfate precipitated is washed, dried, and weighed.

On account of the solubility and hydrolysis of bis( $\beta$ -chloroethyl) sulfide it is quite difficult to estimate with any degree of accuracy small amounts of the sulfide, yet Hopkins (33) introduced a new rapid method whose principle is dependent upon the increase in the hydrogen-ion concentration. By this means a sample can be taken and determination made within five minutes.

#### V. PROPOSED THEORIES TO ACCOUNT FOR VESICANT ACTION

Several theories have been advanced to explain the remarkable vesicant action of bis( $\beta$ -chloroethyl) sulfide. Each fits certain facts but none is entirely satisfactory. The oldest and probably most generally accepted theory is that the vesicant action is due to the formation of free hydrochloric acid within the cell. This was suggested by Marshall (48) and others (43) and substantially supported by Lillie, Clowes, and Chambers (41), but has been opposed by a number of investigators. Warthin and Weller (65) and Arbuckle (4) took exception to it from a pathological standpoint. Peters and Walker (57) showed that the rate of hydrolysis of lipid-soluble substituted diethyl sulfides is not a critical factor.

Flury (20) advanced the theory that the vesicant action was due to the reactivity of the sulfur atom, but when it was found that the sulfoxide containing quadrivalent sulfur was non-vesicant it lost much of its probability. The idea that the unlooked-for reaction of bis( $\beta$ -chloroethyl) sulfide with sodium alcoholate to give the highly toxic divinyl sulfide (5, 29) might have a bearing on the vesicant action was discounted by the later investigations of Bales and Nickelson (6), who found that the yield of divinyl sulfide decreases markedly when the per cent of alcohol drops below 80, and of Lawson and Dawson (38) who obtained no trace of divinyl sulfide from bis( $\beta$ -chloroethyl) sulfide when using secondary sodium phosphate and sodium carbonate in 50 per cent alcohol, which substance gives hydrogen-ion concentrations approaching those in body tissues.

Cashmore and McCombie (11) suggest that the vesicant action of bis( $\beta$ -chloroethyl) sulfide might be due to the product formed by its reaction with the amino acids present in the skin. The results obtained by Lawson and Reid (39) are consistent with the theory that the vesicant action of bis( $\beta$ -chloroethyl) sulfide is due to its reactions with the living cells.

It has also been suggested that the vesicant action is due to the oxidation of the sulfide in the animal system with subsequent hydrolysis of the oxidation product (10).

Arbuckle (4) is of the opinion that the vesicant action is due to the integrity of the entire molecule, that is, the toxicity is maintained only as long as the sulfide remains in the original state.

## VI. PHYSIOLOGICAL PROPERTIES

The symptoms which arise from exposure to bis( $\beta$ -chloroethyl) sulfide consist essentially of local irritation of the upper respiratory tract, eyes, and skin, giving rise to conjunctivitis and superficial necrosis of the cornea, hyperemia, edema, vesication, and later necrosis of the skin, leading to a skin lesion of great chronicity, and congestion and necrosis of the epithelial lining of the pharynx, trachea, and bronchia. In the skin the cells of the epidermis are damaged, following which there is injury to the capillaries with edema, bleb formation, and exudation. The poison may gain entrance to the blood through the injured epithelium and capillaries, and circulatory injury and damage to internal organs result. Injury to the blood itself and particularly the red blood cells consisting of intravascular destruction is evidenced by hemosiderin deposits in the spleen and liver. The severity of the injury to the skin, respiratory organs, and blood corpuscles results in general toxic phenomena manifested by a rapid disappearance of fat from the fat depôts, atrophic changes in the heart and liver, and deposition of waste pigments. In animals exposed to high concentrations of this gas, vomiting and diarrhea, hyperexcitability and convulsions, and

effects on the heart suggest an absorption into the blood stream and a systematic action.

Experiments demonstrate clearly that there is an absorption of dichloroethyl sulfide through the lungs. When this substance is injected either subcutaneously or intravenously into a dog, the following symptoms make their appearance after a latent period: salivation, hyperexcitability, and convulsions, vomiting and bloody diarrhea, a slow and irregular heart, which becomes rapid before death, muscular weakness, and finally, coma and death. Further proof of the absorption through the lungs is furnished by the detection of one of its decomposition products, dihydroxyethyl sulfide, in the urine of the gassed animals (27, 48).

Dichloroethyl sulfide profoundly modifies the leucocyte count of the blood in dogs as well as in rabbits and human beings, and Hektoen and Corper (28) place the sulfide in the same class with the leucotoxic agents: benzene, the radioactive preparation of which thorium X is an example, and the Roentgen ray, all of which have a profound effect on the hemapoietic organs, the leucocytes and specific antibody formation.

Dichloroethyl sulfide is a protoplasmic poison and affects all cells with which it comes in contact (20).

The toxicity of bis( $\beta$ -chloroethyl) sulfide, as determined by Chlopin (12), is 15.3 when compared with chlorine as unity.

## VI. APPLICATIONS

The Mexican bean weevil (*Epilachna corrupta* Muls.) is killed in from three to six hours when 0.01 per cent of the sulfide with a little ether adsorbed on charcoal is dusted over a bean plant. A 4 per cent solution without ether was no more effective; as high as 16 per cent bis( $\beta$ -chloroethyl) sulfide on charcoal did not injure plants. The results with 0.01 per cent sulfide adsorbed on charcoal on the cotton boll weevil (*Anthonomus grandis* Boh.) are also quite promising (47).

Bis( $\beta$ -chloroethyl) sulfide is found to have a direct retarding influence on the development of tuberculosis (15). Adair and Bagg (1) have been able to control tar cancers in mice by the surface application of bis( $\beta$ -chloroethyl) sulfide in absolute alcohol. In human beings it is thought that the sulfide offers another agent for fighting cancer provided the lesion is localized, and Berenblum (8) is able to bring about the inhibition of carcinogenic tar by the addition of a small amount of bis( $\beta$ -chloroethyl) sulfide to the tar.

## REFERENCES

- (1) ADAIR AND BAGG: *Ann. Surgery* **93**, 190-9 (1931).
- (2) ADAMS AND WILLIAMSON: *J. Wash. Acad. Sci.* **9**, 30-5 (1919).
- (3) ADAMS, WILLIAMSON, AND JOHNSON: *J. Am. Chem. Soc.* **41**, 12-42 (1919).
- (4) ARBUCKLE: *Chem. Warfare* **18**, 1078-86 (1932).

- (5) BALES AND NICKELSON: J. Chem. Soc. **121**, 2137-9 (1922).
- (6) BALES AND NICKELSON: J. Chem. Soc. **123**, 2486-9 (1923).
- (7) BENNETT: J. Chem. Soc. **119**, 418-25 (1921).
- (8) BERENBLUM: J. Path. Bact. **34**, 731-46 (1931).
- (9) BOULIN AND SIMON: Compt. rend. **170**, 845-8 (1920).
- (10) CASHMORE: J. Chem. Soc. **123**, 1738-45 (1923).
- (11) CASHMORE AND McCOMBIE: J. Chem. Soc. **123**, 2884-90 (1923).
- (12) CHLOPIN: Z. ges. Schiess-Sprengstoffw. **22**, 227-30 (1927).
- (13) CLARKE: J. Chem. Soc. **101**, 1583-90 (1912).
- (14) CONANT, HARTSHORN, AND RICHARDSON: J. Am. Chem. Soc. **42**, 585-95 (1920).
- (15) CORPER AND RENSCH: J. Infectious Diseases **28**, 286-93 (1920).
- (16) DAVIES: J. Chem. Soc. **117**, 297-308 (1920).
- (17) DAVIES AND OXFORD: J. Chem. Soc. **1931**, 224-36.
- (18) DELÉPINE, FLEURY, AND VILLE: Compt. rend. **172**, 1238-40 (1921).
- (19) DESPRETZ: Ann. chim. phys. **21**, 438 (1822).
- (20) FLURY: Zentr. exptl. Med. **13**, 367-483 (1921).
- (21) GIBSON AND POPE: J. Chem. Soc. **117**, 271-8 (1920).
- (22) GOMBERG: J. Am. Chem. Soc. **41**, 1414-31 (1919).
- (23) GRIGNARD, RIVAT, AND SCATCHARD: Ann. chim. [9] **15**, 1-18 (1921).
- (24) GUTHRIE: Quart. J. Chem. Soc. **12**, 117 (1860).
- (25) HARKINS AND BROWN: J. Am. Chem. Soc. **41**, 499-524 (1919).
- (26) HARKINS AND EWING: J. Am. Chem. Soc. **41**, 1977-80 (1919).
- (27) HEITZMAN: Zentr. exptl. Med. **13**, 484-522 (1921).
- (28) HEKTOEN AND CORPER: J. Infectious Diseases **28**, 279-85 (1920).
- (29) HELFRICH AND REID: J. Am. Chem. Soc. **42**, 1208-32 (1920).
- (30) HERBST: Kolloidchem. Beihefte **23**, 313 (1926).
- (31) HEYCOCK AND MILLS: J. Chem. Soc. **117**, 276 (1920).
- (32) HOLLELY: J. Chem. Soc. **117**, 898-902 (1920).
- (33) HOPKINS: J. Pharmacol. **12**, 393-403 (1919).
- (34) KOSTEVITCH: Z. ges. Schiess-Sprengstoffw. **23**, 365 (1928).
- (35) KRÄMER: Kolloid-Z. **31**, 150-1 (1922).
- (36) KRETOV: J. Russ. Phys. Chem. Soc. **61**, 2345-84 (1929).
- (37) LAWSON AND DAWSON: J. Am. Chem. Soc. **49**, 3119-25 (1927).
- (38) LAWSON AND DAWSON: J. Am. Chem. Soc. **49**, 3125-31 (1927).
- (39) LAWSON AND REID: J. Am. Chem. Soc. **47**, 2821-36 (1925).
- (40) LEWIN: J. prakt. Chem. **118**, 282-6 (1928).
- (41) LILLIE, CLOWES, AND CHAMBERS: J. Pharmacol. **14**, 75-112 (1919).
- (42) LOUGNINE AND DUPONT: Bull. soc. chim. [4] **11**, 904-9 (1912).
- (43) LYNCH, SMITH, AND MARSHALL: J. Pharmacol. **12**, 265-90 (1919).
- (44) MANN AND POPE: J. Chem. Soc. **121**, 594-603 (1922).
- (45) MANN AND POPE: J. Chem. Soc. **121**, 1052-5 (1922).
- (46) MANN, POPE, AND VERNON: J. Chem. Soc. **119**, 634-46 (1921).
- (47) MARCOVITCH: Science **61**, 22 (1925).
- (48) MARSHALL: J. Am. Med. Assoc. **73**, 684-6 (1919).
- (49) MAXIM: Chem. Ztg. **56**, 503 (1932).
- (50) MEYER: Ber. **19**, 3259-66 (1886).
- (51) MUMFORD AND PHILLIPS: J. Chem. Soc. **1928**, 155-62.
- (52) MUMFORD, PHILLIPS, AND BALL: J. Chem. Soc. **1932**, 589-92.
- (53) MYERS AND STEPHENS: J. Soc. Chem. Ind. **39**, 65-6T (1920).
- (54) NIEMANN: Ann. **113**, 288-92 (1860).
- (55) Official History of the Great War, Medical Services, Vol. 1, p. 425 (1923).



- (56) PALMER: J. Chem. Soc. **117**, 276-7 (1920).
- (57) PETERS AND WALKER: Biochem. J. **17**, 260-76 (1923).
- (58) PHILLIPS, DAVIES, AND MUMFORD: J. Chem. Soc. **1929**, 535-49.
- (59) SPICA: Gazz. chim. ital. **49**, II, 299-302 (1919).
- (60) STEINKOPF, HEROLD, AND STAHR: Ber. **53**, 1007-12 (1920).
- (61) THOMPSON, BLACK, AND SOHL: J. Am. Chem. Soc. **43**, 877-9 (1921).
- (62) THOMPSON AND KOPP: J. Ind. Eng. Chem. **12**, 1056-7 (1920).
- (63) THOMPSON AND ODEEN: J. Ind. Eng. Chem. **12**, 1057-62 (1920).
- (64) VEDDER: Medical Aspects of Chemical Warfare, p. 14. Baltimore (1925).
- (65) WARTHIN AND WELLER: The Medical Aspects of Mustard Gas Poisoning. C. V. Mosby Co., St. Louis (1919).
- (66) WILKINSON, NEILSON, AND WYLDE: J. Am. Chem. Soc. **42**, 1377-82 (1920).
- (67) WILKINSON AND WERNLUND: J. Am. Chem. Soc. **42**, 1382-5 (1920).
- (68) WILSON, FULLER, AND SCHUR: J. Am. Chem. Soc. **44**, 2867-78 (1922).
- (69) YABLICK, PERROTT, AND FURMAN: J. Am. Chem. Soc. **42**, 266-74 (1920).





**EMER**

INSTITUTE LIBRARY  
NEW DELHI.

Date of issue.	Date of issue.	Date of issue.
1941	1941	1941
1942	1942	1942
1943	1943	1943
1944	1944	1944
1945	1945	1945
1946	1946	1946
1947	1947	1947
1948	1948	1948
1949	1949	1949
1950	1950	1950
1951	1951	1951
1952	1952	1952
1953	1953	1953
1954	1954	1954
1955	1955	1955
1956	1956	1956
1957	1957	1957
1958	1958	1958
1959	1959	1959
1960	1960	1960
1961	1961	1961
1962	1962	1962
1963	1963	1963
1964	1964	1964
1965	1965	1965
1966	1966	1966
1967	1967	1967
1968	1968	1968
1969	1969	1969
1970	1970	1970
1971	1971	1971
1972	1972	1972
1973	1973	1973
1974	1974	1974
1975	1975	1975
1976	1976	1976
1977	1977	1977
1978	1978	1978
1979	1979	1979
1980	1980	1980
1981	1981	1981
1982	1982	1982
1983	1983	1983
1984	1984	1984
1985	1985	1985
1986	1986	1986
1987	1987	1987
1988	1988	1988
1989	1989	1989
1990	1990	1990
1991	1991	1991
1992	1992	1992
1993	1993	1993
1994	1994	1994
1995	1995	1995
1996	1996	1996
1997	1997	1997
1998	1998	1998
1999	1999	1999
2000	2000	2000
2001	2001	2001
2002	2002	2002
2003	2003	2003
2004	2004	2004
2005	2005	2005
2006	2006	2006
2007	2007	2007
2008	2008	2008
2009	2009	2009
2010	2010	2010
2011	2011	2011
2012	2012	2012
2013	2013	2013
2014	2014	2014
2015	2015	2015
2016	2016	2016
2017	2017	2017
2018	2018	2018
2019	2019	2019
2020	2020	2020
2021	2021	2021
2022	2022	2022
2023	2023	2023
2024	2024	2024
2025	2025	2025
2026	2026	2026
2027	2027	2027
2028	2028	2028
2029	2029	2029
2030	2030	2030
2031	2031	2031
2032	2032	2032
2033	2033	2033
2034	2034	2034
2035	2035	2035
2036	2036	2036
2037	2037	2037
2038	2038	2038
2039	2039	2039
2040	2040	2040
2041	2041	2041
2042	2042	2042
2043	2043	2043
2044	2044	2044
2045	2045	2045
2046	2046	2046
2047	2047	2047
2048	2048	2048
2049	2049	2049
2050	2050	2050
2051	2051	2051
2052	2052	2052
2053	2053	2053
2054	2054	2054
2055	2055	2055
2056	2056	2056
2057	2057	2057
2058	2058	2058
2059	2059	2059
2060	2060	2060
2061	2061	2061
2062	2062	2062
2063	2063	2063